GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 23:00:34 ; Search time 4117 Seconds
(without alignments)
10274.813 Million cell updates/sec

Title: US-10-612-379-1
Perfect score: 873
Sequence: 1 atggcagaagcttaccagat.....ccgatgttaatgttcattaa 873

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	873	100.0	873	3	AY308063	AY308063 Caenorhab
	2	381	43.6	110000	2	CEY105E8_2	Continuation (3 of
c	3	381	43.6	277607	3	CEY105E8A	AL132876 Caenorhab
	4	279	32.0	24174	3	CBRG35I24	AC084558 Caenorhab
	5	238.4	27.3	110000	2	CEY105E8_3	Continuation (4 of
	6	76.4	8.8	1063	6	CQ595182	CQ595182 Sequence
	7	76.4	8.8	1090	6	AR521207	AR521207 Sequence
	8	76.4	8.8	1384	3	AY058721	AY058721 Drosophil
	9	76.4	8.8	1880	6	AR505752	AR505752 Sequence
c	10	51.2	5.9	2000	6	AX655393	AX655393 Sequence
	11	48	5.5	877	9	CR533501	CR533501 Homo sapi
	12	48	5.5	944	9	BC012444	BC012444 Homo sapi
	13	48	5.5	999	9	AF097330	AF097330 Homo sapi
	14	48	5.5	2238	6	AX833868	AX833868 Sequence
	15	48	5.5	2238	9	AK095959	AK095959 Homo sapi
c	16	47.6	5.5	28057	3	CEF26H11	Z81515 Caenorhabdi
c	17	47.4	5.4	7218	6	I66494	I66494 Sequence 14
	18	47.2	5.4	450	6	CQ721554	CQ721554 Sequence
	19	46.4	5.3	819	6	AR149778	AR149778 Sequence
	20	46.4	5.3	819	6	BD190878	BD190878 Human ani
	21	46.4	5.3	819	6	AR271589	AR271589 Sequence
	22	45.4	5.2	1397	5	BC076239	BC076239 Danio rer
	23	44.8	5.1	1633	9	AF109196	AF109196 Homo sapi
c	24	44.6	5.1	95660	2	AC131181	AC131181 Homo sapi
c	25	44.6	5.1	107755	9	AC116628	AC116628 Homo sapi
c	26	44.6	5.1	156817	9	AC009406	AC009406 Homo sapi
	27	43.4	5.0	842	6	CQ716785	CQ716785 Sequence
	28	43.4	5.0	1229	6	BD094065	BD094065 Shear str
	29	43.4	5.0	1229	9	HSP64BCCP	Y12696 Homo sapien
	30	43.4	5.0	1604	9	BC022305	BC022305 Homo sapi
	31	43.4	5.0	1640	9	BC005367	BC005367 Homo sapi
	32	43.4	5.0	1652	6	AR339324	AR339324 Sequence
	33	43.4	5.0	1722	6	CQ491450	CQ491450 Sequence
	34	43.4	5.0	1722	6	CQ497331	CQ497331 Sequence
	35	43	4.9	10560	6	CQ595181	CQ595181 Sequence
	36	43	4.9	44319	2	AC013191	AC013191 Drosophil
c	37	43	4.9	92969	2	AC010844	AC010844 Drosophil
	38	43	4.9	142372	9	AC132152	AC132152 Homo sapi
c	39	43	4.9	168384	3	AC011069	AC011069 Drosophil
	40	43	4.9	209320	2	AC060776	AC060776 Homo sapi
c	41	43	4.9	217422	9	AC108721	AC108721 Homo sapi
c	42	43	4.9	307761	3	AE003493	AE003493 Drosophil
	43	42.8	4.9	447	9	AF426169	AF426169 Homo sapi
	44	42.8	4.9	511	4	AF109198	AF109198 Bos tauru
	45	42.8	4.9	859	6	AX921765	AX921765 Sequence

ALIGNMENTS

RESULT 1
AY308063
LOCUS AY308063 873 bp mRNA linear INV 21-NOV-2003
DEFINITION Caenorhabditis elegans EXC-4 (exc-4) mRNA, complete cds.
ACCESSION AY308063
VERSION AY308063.1 GI:34559858
KEYWORDS .
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 873)
AUTHORS Berry,K.L., Bulow,H.E., Hall,D.H. and Hobert,O.
TITLE A Caenorhabditis elegans CLIC-like protein required for intracellular tube formation and maintenance
JOURNAL Science (2003) In press
REFERENCE 2 (bases 1 to 873)
AUTHORS Berry,K. and Hobert,O.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2003) Biochemistry, Columbia University, 701 W.168th St., New York, NY 10032, USA
FEATURES Location/Qualifiers
source 1. .873
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
gene 1. .873
/gene="exc-4"
CDS 1. .873
/gene="exc-4"
/note="ion channel"
/codon_start=1
/product="EXC-4"
/protein_id="AAQ75554.1"
/db_xref="GI:34559859"
/translation="MAEAYQIQSNGDPQSKPLLELYVKASGIDARRIGADLFCQEFWM
ELYALYEIGVARVEVKTVNVNSEAFKKNFLGAQPPIMIEEEKELTYTDNREIEGRIFH
LAKEFNVPLFEKDPSAEKRIENLYRNFKLFLRAKVEFDKGKKEPSRVEDLPAQIKVHY
NRVCEQLSNIDQLLSERKSRYLLGNSMTEYDCELMPRLHHIRIIGLSLLGFDIPHNFT
HLWAYILTAYRTAAFIESCPADQDIIHHYKEQMNLFTNQRETLQSPTKTHTIPEKVLS
DIRVKGLAPDVNVH"
ORIGIN

```
  Query Match           100.0%;  Score 873;  DB 3;  Length 873;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-219;
  Matches  873;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCAGAAGCTTACCAGATCCAATCAAACGGAGATCCCCAATCAAAACCTCTTCTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGCAGAAGCTTACCAGATCCAATCAAACGGAGATCCCCAATCAAAACCTCTTCTCGAG 60

Qy         61 CTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAG 120

Qy        121 GAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAG 180

Qy        181 ACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATT 240

Qy        241 ATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATC 300

Qy        301 TTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAG 360

Qy        361 AGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGAT 420

Qy        421 AAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTAC 480

Qy        481 AATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGA 540

Qy        541 TATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCAT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCAT 600

Qy        601 ATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTC 660

Qy        661 TGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGAT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGAT 720

Qy        721 CAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACC 780

Qy        781 CTCCAATCGCCAACAAAAACGCACACAATTCCGGAAAAAGTGCTATCGGATATTCGTGTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTCCAATCGCCAACAAAAACGCACACAATTCCGGAAAAAGTGCTATCGGATATTCGTGTT 840

Qy        841 AAAGGACTTGCTCCCGATGTTAATGTTCATTAA 873
              ||||||||||||||||||||||||||||||||
Db        841 AAAGGACTTGCTCCCGATGTTAATGTTCATTAA 873


RESULT 2
CEY105E8_2
WPCOMMENT
  Sequence split into 7 fragments  LOCUS CEY105E8 Accession AL022594
     Fragment Name      Begin     End
     CEY105E8_0             1    110000
     CEY105E8_1        100001    210000
     CEY105E8_2        200001    310000
     CEY105E8_3        300001    410000
     CEY105E8_4        400001    510000
     CEY105E8_5        500001    610000
     CEY105E8_6        600001    702117
Continuation (3 of 7) of CEY105E8 from base 200001 (AL022594 Caenorhabditis elegans chro

  Query Match           43.6%;  Score 381;  DB 2;  Length 110000;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-89;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCAGAAGCTTACCAGATCCAATCAAACGGAGATCCCCAATCAAAACCTCTTCTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91763 ATGGCAGAAGCTTACCAGATCCAATCAAACGGAGATCCCCAATCAAAACCTCTTCTCGAG 91822

Qy         61 CTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91823 CTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAG 91882

Qy        121 GAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91883 GAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAG 91942

Qy        181 ACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      91943 ACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATT 92002

Qy        241 ATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92003 ATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATC 92062

Qy        301 TTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92063 TTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAG 92122

Qy        361 AGAATAGAGAACTTGTACAGG 381
              |||||||||||||||||||||
Db      92123 AGAATAGAGAACTTGTACAGG 92143


RESULT 3
CEY105E8A/c
LOCUS       CEY105E8A              277607 bp    DNA     linear   INV 12-OCT-2004
DEFINITION  Caenorhabditis elegans YAC Y105E8A, complete sequence.
ACCESSION   AL132876
VERSION     AL132876.4  GI:18376550
KEYWORDS    HTG.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1
  AUTHORS   none.
  TITLE     Genome sequence of the nematode C. elegans: a platform for
            investigating biology. The C. elegans Sequencing Consortium
  JOURNAL   Science 282 (5396), 2012-2018 (1998)
  MEDLINE   99069613
   PUBMED   9851916
  REMARK    The C.elegans Sequencing Consortium.
REFERENCE   2  (bases 1 to 277607)
  AUTHORS   Sulston,J.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-OCT-2004) Nematode Sequencing Project, Sanger
            Institute, Hinxton, Cambridge CB10 1SA, England and Department of
            Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
            jes@sanger.ac.uk or rw@nematode.wustl.edu E-mail: worm@sanger.ac.uk
COMMENT     On Jan 25, 2002 this sequence version replaced gi:15130764.
            Coding sequences below are predicted from computer analysis, using
            predictions from Genefinder (P. Green, U. Washington), and other
            available information.
            Current sequence finishing criteria for the C. elegans genome
            sequencing consortium are that all bases are either sequenced
            unambiguously on both strands, or on a single strand with both a
```

```
            dye primer and dye terminator reaction, from distinct subclones.
            Exceptions are indicated by an explicit note.
            IMPORTANT:  This sequence is NOT necessarily the entire insert of
            the specified clone.  It may be shorter because we only sequence
            overlapping sections once, or longer because we arrange for a small
            overlap between neighbouring submissions.
            For a graphical representation of this sequence and its analysis
            see:- http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?
            name=Y105E8A
            For a graphical representation of this sequence and its analysis
            see:- http:
            IMPORTANT: This sequence is not the entire insert of clone Y105E8A.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we arrange for a small overlap between
            neighbouring submissions.
            The true left end of clone Y54E5 is at 356131 in this sequence. The
            true right end of clone F49B2 is at 104 in this sequence. The start
            of this sequence (1. .104) overlaps with the end of sequence Z81543.
            The end of this sequence (277504. .277607) overlaps with the start
            of sequence AL132877.
FEATURES             Location/Qualifiers
     source          1. .277607
                     /organism="Caenorhabditis elegans"
                     /mol_type="genomic DNA"
                     /strain="Bristol N2"
                     /db_xref="taxon:6239"
                     /chromosome="I"
                     /clone="Y105E8A"
     gene            complement(join(8762. .8781,8884. .8953,9025. .9099,
                     10412. .10645))
                     /gene="Y105E8A.1"
     CDS             complement(join(8762. .8781,8884. .8953,9025. .9099,
                     10412. .10645))
                     /gene="Y105E8A.1"
                     /standard_name="Y105E8A.1"
                     /note="contains similarity to Pfam domain PF04908
                     (SH3-binding, glutamic acid-rich protein)"
                     /codon_start=1
                     /product="Hypothetical protein Y105E8A.1"
                     /protein_id="CAB60842.2"
                     /db_xref="GI:38422775"
                     /db_xref="UniProt/TrEMBL:Q9NF23"
                     /translation="MAPLKVYVASATANPETKYRVQRTLMILDGLGIPFDSIDITKPE
                     HAEQRRFMRENASKKGPNGAVLPPQFFYEDEYLGDYEDFDTSVEADTITEFLRLLPDA
                     IGNRIRVNDANKCVAALFSKTQGAGVKCNH"
     gene            complement(join(14577. .14881,16799. .16871,18301. .18553,
                     21844. .22457,24788. .25147))
                     /gene="Y105E8A.2"
     CDS             complement(join(14577. .14881,16799. .16871,18301. .18553,
                     21844. .22457,24788. .25147))
                     /gene="Y105E8A.2"
                     /standard_name="Y105E8A.2"
                     /note="contains similarity to Interpro domain IPR009011
                     ()"
                     /codon_start=1
                     /product="Hypothetical protein Y105E8A.2"
                     /protein_id="CAB60843.3"
                     /db_xref="GI:31442020"
                     /db_xref="UniProt/TrEMBL:Q9NF22"
                     /translation="MVSTGLFRPTKAIPLLFLLFSLYKTEKIDDSIHYYVSFTSDSVT
                     IKDPPTLETGNTELSDDYVRIMSKNNERFLCKLPQIDVEVKKNVEYSGATAGKLLETM
                     LYKDKMCSYLIDVYWTYQVCHGRYVIQYHEDKMLTGQVSRTEFYLGNFDSALTASTNE
                     QVKPATRRIENEDYPYYSVSYNHGTSCDVTGGKPRTTDVVYICVEKVQHKILSVTEIS
                     SCHYEIVIMTDLLCKHPEYQLSEKKDHKIVCWNEDAKNENEAKPQTLQRLDDFHDSTF
                     KREYTVFTDQDPQYSENTEEEDGGLIDEKDFEKLKETRQAGYILGGVKRFNTKESLAN
                     NPAVVHHTVNRILTGEDCIVGGTGWWKYEFCYGKHVIQFHEDANGQRSDILLGVFDEV
                     VHKEWVKLDRHARGAIQGNNQIDQISQIYAKGDICDETGAHNVEVRIRCATADHSALS
                     FSMHLTEPRTCQYVLTIDSERFCEPLQFADDYGLIELTQVASSSQSAAAAAAAPVELS
                     IDEEEEDEAHEQSDEDDHDVKIHEDL"
     gene            join(25501. .25562,25607. .25729,25795. .26120,26746. .26939,
                     27768. .28086,29660. .30045,31473. .31976,33263. .33496,
                     34152. .34244)
                     /gene="Y105E8A.3"
     CDS             join(25501. .25562,25607. .25729,25795. .26120,26746. .26939,
                     27768. .28086,29660. .30045,31473. .31976,33263. .33496,
                     34152. .34244)
                     /gene="Y105E8A.3"
                     /standard_name="Y105E8A.3"
                     /note="contains similarity to Pfam domain PF01545 (Cation
                     efflux family)"
                     /codon_start=1
                     /product="Hypothetical protein Y105E8A.3"
                     /protein_id="CAD21657.2"
                     /db_xref="GI:24817580"
                     /db_xref="GOA:Q8WQB6"
                     /db_xref="UniProt/TrEMBL:Q8WQB6"
                     /translation="MLAHDFVKKDDRNFTPTQSSRSASYFVVLVLTKILRCIGVFFVD
                     ILAKSSHIVNLLWLCKIICTCILFPLQKPLSAGKSLKRGIMPMVLKLSLMNSVIEVFW
                     FYGITLCGPLRSVLVFELSSTVLLSAVVSFFKGGNMSTPSKTRGFFLLIVGFVALFLM
                     DRDGTIEDPHHTTHAHHSGLNHLFYHLIGLFGMADHKGGLCILLISLLLRTGYETSFR
                     HLAVEVGGAKRLHTLVTAGSAVCLTPVAVLSWALSSSSGASIGYFHYAALLVIIAVFV
                     FVLDFYAESICFQHVADPVMAAARWSSVTMFSCAFGLAYLWYGSQNLGDHALTGGVSI
                     TVVCFILASISLTHSNQPKHRGGHFVGISNTGLPLFTYGEAFLQRTSKSLMLFMKETL
                     NEILMNNDSRRIFWFLCVNLGFCGVEFLYGFWTNSLGLISDGFHMLFDCSALVMGLVA
                     SVMARWPPTRHFTFGFGRVEILSGFINALFLCVIALFILIEALERLFDPPNINTDRLL
                     FVAISGLIVNLFGMYSLGEHGHSHGGGSSHGHSHGGGGSHGHSHGGGGNANMQGVFLH
                     VLADTLGSVFVIISTLLIQWFGWVWVDPLCSLILSLLIIGSVYPLLVSSISTLLQDVP
                     EEEEFEYHINEILEIEHVESYSNAHMWQHKSDINVASVHVQVKEEANAQMIRHRVSNI
                     LKSTGATHSTVQVEKKTFAHRIQQVCPGYKAGYTVVRGSVIREKKHDHSHHHGHSHDD
                     HSGHHH"
     gene            complement(join(34458. .34607,35321. .35699,36371. .36612))
                     /gene="Y105E8A.4"
     CDS             complement(join(34458. .34607,35321. .35699,36371. .36612))
                     /gene="Y105E8A.4"
                     /standard_name="Y105E8A.4"
                     /note="contains similarity to Pfam domain PF00378
                     (Enoyl-CoA hydratase/isomerase)"
                     /codon_start=1
                     /product="Hypothetical protein Y105E8A.4"
                     /protein_id="CAC48118.1"
                     /db_xref="GI:15130766"
                     /db_xref="GOA:Q9NEZ8"
                     /db_xref="UniProt/TrEMBL:Q9NEZ8"
                     /translation="MITLSRTGAAENVALITLNRPSALNALCRELMLELSENLLKVEK
                     DQSYHVIVLTGSEKAFAAGADIKEMAKLEFADVFENDYFTNWDTLSHITKPVIAAVNG
                     FALGGGTELALMCDIVYAGENAIFGQPEITIGTIPGLGGTQRWPRYVSKSVAMEICLS
                     GDRLGAQEAKEDGLVSKVFPVQQLVGEAVLLADRIAKNSPLIVKTVKRSLNSAYQTSL
                     NQGLEMEKQLFQSTFATNDRREGMSAFAEKRAPKWTSS"
     gene            complement(join(38162. .38414,39163. .39338,39389. .39748,
                     40251. .40334,40489. .40628,40809. .40974,41025. .41246,
                     41384. .41458,42603. .42857))
                     /gene="Y105E8A.5"
     CDS             complement(join(38162. .38414,39163. .39338,39389. .39748,
                     40251. .40334,40489. .40628,40809. .40974,41025. .41246,
                     41384. .41458,42603. .42857))
                     /gene="Y105E8A.5"
                     /standard_name="Y105E8A.5"
                     /note="C. elegans BBS-1 protein; contains similarity to
                     Red clover mottle virus Genome polyprotein B [Contains:
                     Protease cofactor; Membrane bindingsprotein; VPG; Protease
                     (EC 3.4.22.-); RNA polymerase (EC 2.7.7.48)].;
                     SW:VGNB_RCMV"
                     /codon_start=1
                     /product="Hypothetical protein Y105E8A.5"
                     /protein_id="CAC48119.1"
                     /db_xref="GI:15130767"
                     /db_xref="UniProt/TrEMBL:Q9NEZ7"
                     /translation="MAKPVNVNQSKWTVPVLLKECEIYCPSTCVAFGPILSDNDSKLI
                     IAHGGHRGVNMKLKVFQQLEQLSESSLADMPTALVHFINDLSSIPSIAVAAGPSLLIY
                     KNLKPFYKFTVPSSAINPTESEAWKAVVNKKINGDTLLTVLKRLEDDVAFSKLTPISQ
                     TYLRADKETQVVLVEHYGTKIANSATITCIAKLTKSTAEPLDILVIGTEHCEIFLIDS
                     QAFTILETIKIGSVPVNICAYGTYDVDYRLFVQTRASLIFCMKRGEADYQPIVISQSM
                     ITSMVLVNKMIVYTTVENLIHFASFRGKKMNTVKCPSKIKMLEPFIYPLKQLAAVIAV
                     FDKEIRMYNEHYLLDTVQYEKPLAWVKYGCYGREDSTLVVAFKDGSIAIQIFRRKANF
                     DTKLDYNQVPQAHALKLQIPKKTKVFIDLTQREVQLGNRIHKVYQKNLFDVKYRLAAS
                     YLELTSSASATVSTTTVLPVEISVDIHGFGPTFRMTIHLLSSSKQNLYDMHLSIISDP
```

```
                ELYDFDTPLIPVHLLASGQSYSFTTLLYCKDPEKAANCDVRALLVHAKRATPIVTAVI
                KMPFSEFPLD"
     gene       complement(join(48295. .48698,48749. .48937,49598. .49745,
                49796. .49884,51355. .51577))
                /gene="Y105E8A.6"
     CDS        complement(join(48295. .48698,48749. .48937,49598. .49745,
                49796. .49884,51355. .51577))
                /gene="Y105E8A.6"
                /standard_name="Y105E8A.6"
                /note="C. elegans UNC-95 protein; contains similarity to
                Pfam domain PF00412 (LIM domain containing proteins)"
                /codon_start=1
                /product="Hypothetical protein Y105E8A.6"
                /protein_id="CAC48121.2"
                /db_xref="GI:24817579"
                /db_xref="UniProt/TrEMBL:Q9NEZ5"
                /translation="MTISPQPSHQQFESYQWTTESRSSQQRHGTGTPSQDGRLSAIPD
                PVERHVARWRSESRNSNKDKVFRNDEEFSQQDEIVNGTLTALKNDVEQTTEIIRRKQE
                QMRMERRQFQTEMEVNGRISIDPTDDWLAARLKAVSSDDMNQQLVKLKQDQRQNAVTD
                TLAALVYDVNATTEVLRRGQRGRDGEDGNKKKKEEIEYTLRLTPAPEEQIPQRPKIPE
                DDNMETDDYSRQYGVQMSEETDSLRRRRARSTTPRRTLHISGSPPPPAAAVCAYCSEE
                IDGAILTALAPNSERAQKFHTYHFMCTYCQKALNMHGTYREHDLKPYCHDCFYKLYNG
                LQYAPDDHQASIEKLI"

  Query Match          43.6%;  Score 381;  DB 3;  Length 277607;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-89;
  Matches  381;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGCAGAAGCTTACCAGATCCAATCAAACGGAGATCCCCAATCAAAACCTCTTCTCGAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185555 ATGGCAGAAGCTTACCAGATCCAATCAAACGGAGATCCCCAATCAAAACCTCTTCTCGAG 185496

Qy         61 CTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185495 CTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAG 185436

Qy        121 GAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185435 GAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAG 185376

Qy        181 ACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185375 ACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATT 185316

Qy        241 ATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185315 ATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATC 185256

Qy        301 TTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     185255 TTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAG 185196

Qy        361 AGAATAGAGAACTTGTACAGG 381
              |||||||||||||||||||||
Db     185195 AGAATAGAGAACTTGTACAGG 185175


RESULT 4
CBRG35I24
LOCUS       CBRG35I24              24174 bp    DNA     linear   INV 04-NOV-2000
DEFINITION  Caenorhabditis briggsae cosmid G35I24, complete sequence.
ACCESSION   AC084558
VERSION     AC084558.1  GI:11095008
KEYWORDS    HTG.
SOURCE      Caenorhabditis briggsae
  ORGANISM  Caenorhabditis briggsae
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 24174)
  AUTHORS   Washington University Genome Sequencing Center.
  TITLE     The C. briggsae Genome Sequencing Project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 24174)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     Submitted by:
                    Genome Sequencing Center
                    Department of Genetics, Washington University,
                    St. Louis, MO 63110, USA
                    e-mail: jspieth@watson.wustl.edu


            NOTICE:  This sequence may not be the entire insert of this clone.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we provide a small overlap between
            neighboring submissions.
FEATURES            Location/Qualifiers
     source         1. .24174
                    /organism="Caenorhabditis briggsae"
                    /mol_type="genomic DNA"
                    /strain="GujArat G16"
                    /db_xref="taxon:6238"
                    /clone="G35I24"
ORIGIN

  Query Match          32.0%;  Score 279;  DB 3;  Length 24174;
  Best Local Similarity  80.3%;  Pred. No. 1.1e-62;
  Matches  327;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy        348 ATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCAAA 407
              | ||||  |   |   |  ||  || | |  ||||||||||| ||||||| || ||
Db      23067 ACCCGCAAATTGAAATTCATTGATTTTTTTCTAGAACTTCAAACTCTTCCTGCGCGCCAA 23126

Qy        408 AGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCACAGAT 467
              ||| || || || |||||||| |||||||||||||||| || ||  | ||||||||||
Db      23127 AGTCGAATTTGACAAGGGAAAGAAGGAGCCATCGAGAGTGGAGGACTTGCCAGCACAGAT 23186

Qy        468 TAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTATCCGA 527
              |||||  || |||||  || |||||||||| || ||||||||||||| || |||||||
Db      23187 CAAAGTGCATTACAACCGTGTCTGTGAGCAGCTCTCCAATATTGATCAGCTGTTATCCGA 23246

Qy        528 GAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCC 587
              ||||||  | || ||  | |||||||||||||||| || |||||||||||| |||||||
Db      23247 GAGAAAGACACGCTACTTGCTTGGAAACAGTATGACAGAGTATGACTGTGAATTGATGCC 23306

Qy        588 ACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAA 647
              |||||||||||| ||||| || ||||| ||||| ||||||||||||||| || ||||| ||
Db      23307 ACGTCTTCATCACATTCGTATCATTGGGTTGTCGCTTCTTGGATTCGACATCCCACACAA 23366

Qy        648 TTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAG 707
              ||||||||||||||||    || || ||||| || |||||||| || || || ||||||||
Db      23367 TTTCACTCATCTCTGGAACTACATTCTCACCGCTTACCGTACCGCGGCGTTCATTGAGAG 23426

Qy        708 TTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGA 754
              ||||| |||||||||| || |||||||| ||||| | |  || ||
Db      23427 CTGTCCAGCCGATCAGGATATAATTCATCATTATAAGGTAGGAAAGA 23473


RESULT 5
CEY105E8_3
WPCOMMENT
  Sequence split into 7 fragments  LOCUS CEY105E8 Accession AL022594
     Fragment Name      Begin     End
     CEY105E8_0             1     110000
     CEY105E8_1        100001     210000
     CEY105E8_2        200001     310000
     CEY105E8_3        300001     410000
     CEY105E8_4        400001     510000
     CEY105E8_5        500001     610000
     CEY105E8_6        600001     702117
Continuation (4 of 7) of CEY105E8 from base 300001 (AL022594 Caenorhabditis elegans chro

  Query Match          27.3%;  Score 238.4;  DB 2;  Length 110000;
  Best Local Similarity  93.9%;  Pred. No. 5.3e-52;
```

```
  Matches  248;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy        503 CCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGA 562
              ||| | |    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15738 CCATTTTATTGCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGA 15797

Qy        563 CTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCAC 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15798 CTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCAC 15857

Qy        623 TTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCAT 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15858 TTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCAT 15917

Qy        683 ACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATA 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      15918 ACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATA 15977

Qy        743 AAGAACAAATGAATCTGTTCACAA 766
              |||     ||||| |||   || |
Db      15978 AAGTGAGGATGAAGCTGCAAACTA 16001


RESULT 6
CQ595182
LOCUS       CQ595182               1063 bp    DNA     linear   PAT 02-FEB-2004
DEFINITION  Sequence 22940 from Patent WO0171042.
ACCESSION   CQ595182
VERSION     CQ595182.1  GI:41652101
KEYWORDS    .
SOURCE      Drosophila sp.
  ORGANISM  Drosophila sp.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Venter,J.C., Adams,M., Li,P.W. and Myers,E.W.
  TITLE     Detection kits, such as nucleic acid arrays, for detecting the
            expression of 10,000 or more Drosophila genes and uses thereof
  JOURNAL   Patent: WO 0171042-A 22940 27-SEP-2001;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .1063
                     /organism="Drosophila sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:7242"
ORIGIN

  Query Match             8.8%;  Score 76.4;  DB 6;  Length 1063;
  Best Local Similarity   52.5%;  Pred. No. 4.1e-09;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    || |
Db        727 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 786

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              || |||      |  |||||||| |||||||| || || || || || ||  |    ||
Db        787 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 846

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||  || | |  || || || ||     | || ||    || |||   || ||
Db        847 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 906

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                |||     |  || || |   |    || || ||||| |||||||||||  |
Db        907 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 966

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||   |||  || |||    |  || | |  ||    ||  |  |||| ||
Db        967 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 1026

Qy        795 AAAAACGCACACAATTCC 812
```

```
                ||       |||||
Db      1027 GTTTACCACATACATTCC 1044


RESULT 7
AR521207
LOCUS       AR521207               1090 bp    DNA     linear   PAT 22-SEP-2004
DEFINITION  Sequence 26167 from patent US 6703491.
ACCESSION   AR521207
VERSION     AR521207.1  GI:52456682
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1090)
  AUTHORS   Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L.,
            Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
  TITLE     Drosophila sequences
  JOURNAL   Patent: US 6703491-A 26167 09-MAR-2004;
FEATURES             Location/Qualifiers
     source          1. .1090
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.8%;  Score 76.4;  DB 6;  Length 1090;
  Best Local Similarity   52.5%;  Pred. No. 4.1e-09;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    || |
Db         28 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 87

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              || |||      |  |||||||| |||||||| || || || || || ||  |    ||
Db         88 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 147

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||  || | |  || || || ||     | || ||    || |||   || ||
Db        148 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 207

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                |||     |  || || |   |    || || ||||| |||||||||||  |
Db        208 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 267

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||   |||  || |||    |  || | |  ||    ||  |  |||| ||
Db        268 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 327

Qy        795 AAAAACGCACACAATTCC 812
                ||       |||||
Db        328 GTTTACCACATACATTCC 345


RESULT 8
AY058721
LOCUS       AY058721               1384 bp    mRNA    linear   INV 16-OCT-2001
DEFINITION  Drosophila melanogaster LD46306 full length cDNA.
ACCESSION   AY058721
VERSION     AY058721.1  GI:16185714
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1384)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (08-OCT-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone.  However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .1384
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="12B9-12B9"
     gene            1. .1384
                     /gene="CG10997"
                     /note="alignment with genomic scaffold AE003493"
                     /db_xref="FLYBASE:FBgn0030529"
     CDS             317. .1099
                     /gene="CG10997"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="LD46306p"
                     /protein_id="AAL13950.1"
                     /db_xref="GI:16185715"
                     /db_xref="FLYBASE:FBgn0030529"
                     /translation="MSEVESQQSQETNGSSKFDVPEIELIIKASTIDGRRKGACLFCQ
                     EYFMDLYLLAELKTISLKVTTVDMQKPPPDFRTNFEATHPPILIDNGLAILENEKIER
                     HIMKNIPGGYNLFVQDKEVATLIENLYVKLKLMLVKKDEAKNNALLSHLRKINDHLSA
                     RNTRFLTGDTMCCFDCELMPRLQHIRVAGKYFVDFEIPTHLTALWRYMYHMYQLDAFT
                     QSCPADQDIINHYKLQQSLKMKKHEELETPTFTTYIPIDISE"
ORIGIN

  Query Match             8.8%;  Score 76.4;  DB 3;  Length 1384;
  Best Local Similarity   52.5%;  Pred. No. 4e-09;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    ||  |
Db        763 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 822

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              ||  |||      |  |||||||| |||||||| || || || || || ||  |   ||
Db        823 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 882

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||  || | |  || || || ||     | || ||    || |||   || ||
Db        883 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 942

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                ||||     |  || || |   |    || || ||||| |||||||||||  |
Db        943 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 1002

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||    |||  || |||    |  ||  | |  ||    ||  |  |||| ||
Db       1003 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 1062

Qy        795 AAAAACGCACACAATTCC 812
```

```
                ||      |||||
Db       1063 GTTTACCACATACATTCC 1080


RESULT 9
AR505752
LOCUS       AR505752               1880 bp    DNA     linear   PAT 22-SEP-2004
DEFINITION  Sequence 10712 from patent US 6703491.
ACCESSION   AR505752
VERSION     AR505752.1  GI:52441227
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1880)
  AUTHORS   Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L.,
            Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
  TITLE     Drosophila sequences
  JOURNAL   Patent: US 6703491-A 10712 09-MAR-2004;
FEATURES             Location/Qualifiers
     source          1. .1880
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

  Query Match             8.8%;  Score 76.4;  DB 6;  Length 1880;
  Best Local Similarity   52.5%;  Pred. No. 4e-09;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    ||  |
Db        818 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 877

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              ||  |||      |  |||||||| |||||||| || || || || || ||  |   ||
Db        878 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 937

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||  || | |  || || || ||     | || ||    || |||   || ||
Db        938 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 997

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                ||||     |  || || |   |    || || ||||| |||||||||||  |
Db        998 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 1057

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||    |||  || |||    |  ||  | |  ||    ||  |  |||| ||
Db       1058 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 1117

Qy        795 AAAAACGCACACAATTCC 812
                ||      |||||
Db       1118 GTTTACCACATACATTCC 1135


RESULT 10
AX655393/c
LOCUS       AX655393               2000 bp    DNA     linear   PAT 22-MAR-2003
DEFINITION  Sequence 5263 from Patent WO03000898.
ACCESSION   AX655393
VERSION     AX655393.1  GI:29158207
KEYWORDS    .
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,
            Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.
  TITLE     Plant genes involved in defense against pathogens
  JOURNAL   Patent: WO 03000898-A 5263 03-JAN-2003;
            Syngenta Participations AG (CH)
FEATURES             Location/Qualifiers
```

```
     source          1. .2000
                     /organism="Oryza sativa"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4530"
ORIGIN

  Query Match              5.9%;  Score 51.2;  DB 6;  Length 2000;
  Best Local Similarity   10.4%;  Pred. No. 0.018;
  Matches   70;  Conservative  294;  Mismatches  302;  Indels    4;  Gaps    1;

Qy         75 GTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAGGAATTCTGGATGGA 134
Db       1030 GTGTGTAGGTGCTKMWRTYTRSMSWTYAMWKKYTKYMTAYSSTWKWYWAYKWRAYAWSRS 971

Qy        135 GTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAGACTGTCAACGTGAA 194
Db        970 RKTWWCTGGKRMATYCGTKMWAAGRWRWRMAWCWYCCMWKWWKMTSCMWWKYWRTWSCWY 911

Qy        195 TTCTGAAGCATTTAAGAA----GAACTTTCTCGGAGCACAACCACCGATTATGATTGAAG 250
Db        910 TMWWGAMRYAYYAMRRRRWTYKWSWRRMYWTMTKWAWTWMTCMCMAKWYMATGWATWMWW 851

Qy        251 AGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATCTTTCATTTGG 310
Db        850 RYTMYTYCYAMTCAKCKYKMAMTKWWTTWACAWRATSWRWRAMAGMRWKRYKMKRAYWWR 791

Qy        311 CAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGA 370
Db        790 WWRCWKAGWARWMKSRYRWKWKKYATRYYWKMWAMTWWWSWRRWKSYRMWSGMGRMRWSA 731

Qy        371 ACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAA 430
Db        730 WRYCSRMKCAKTKYASSARWTKRAKRSYRYRRRWYWKRKGWTYRYRYWRSCRMTRARMSK 671

Qy        431 AGGAGCCATCGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCT 490
Db        670 RRKWAGASMKSCWMYWRGARSMWYSKYSCSAKCCKKTRYMTSSYMSTGMYGMYSSYKSMS 611

Qy        491 GTGAGCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTG 550
Db        610 WTSKMSYMGKMTCTMYTSMKGSTRRSKMGRWSGMSRMYMRWWKKMRKRKYMRYMKWKCTW 551

Qy        551 GAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTA 610
Db        550 RRCMCYRWGYTMYTTSRSRMMYTGRYKARYTSKRRYMWYKYRKYCWYYYYGMYMKCSYMM 491

Qy        611 TTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATA 670
Db        490 RYGYCKACKKCCYAMCWKAAYSGMMMYWYRKYSKWMRMSTKYMWSMWYKKCRSMKYGAKG 431

Qy        671 TCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTA 730
Db        430 CYGCKMWTYCSYGYMKWYTYMGSYKYSRCYKYMRMYMYKGWMYMMYYSAYSSMMTWYYYY 371

Qy        731 TTCATCACTA 740
Db        370 AKYWKYWYKR 361


RESULT 11
CR533501
LOCUS       CR533501                877 bp    mRNA    linear   PRI 22-JUN-2004
DEFINITION  Homo sapiens full open reading frame cDNA clone RZPDo834E0417D for
            gene CLIC4, chloride intracellular channel 4; complete cds, incl.
            stopcodon.
ACCESSION   CR533501
VERSION     CR533501.1  GI:49065427
KEYWORDS    Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 762)
  AUTHORS   Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.
  TITLE     Cloning of human full open reading frames in Gateway(TM) system
            entry vector (pDONR201)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 762)
  AUTHORS   Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
            Germany
COMMENT     RZPD; RZPDo834E0417D, ORFNo 2835
            www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDo834E0417D RZPDLIB;
            Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
            834
            www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
            www.rzpd.de/products/orfclones/
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            contact RZPD (customer.service@rzpd.de) for further information.
            This CDS clone is a part of a collection of human full ORF clones
            generated by RZPD.
            This CDS has been cloned incl. stopcodon.
            The CDS has been inserted into pDONR201 via a BP Clonase(TM)
            reaction. Additional sequence has been added in front of the start
            codon: att. .AAAAAA GCA GGC (ATG).
            After the stop codon 3' UTR sequence is present in front of the 3'
            att site (ACCCAGCTTTCTT).
            Compared to the reference sequence NM_013943 (gi7330334) we did not
            find any amino acid exchanges.
            Clone distribution: http://www.rzpd.de/products/orfclones/.
FEATURES             Location/Qualifiers
     source          1. .877
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDo834E0417D"
                     /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
                     /lab_host="DH10B"
                     /note="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
     gene            1. .877
                     /gene="CLIC4"
     CDS             1. .762
                     /gene="CLIC4"
                     /codon_start=1
                     /protein_id="CAG38532.1"
                     /db_xref="GI:49065428"
                     /translation="MALSMPLNGLKEEDKEPLIELFVKAGSDGESIGNCPFSQRLFMI
                     LWLKGVVFSVTTVDLKRKPADLQNLAPGTHPPFITFNSEVKTDVNKIEEFLEEVLCPP
                     KYLKLSPKHPESNTAGMDIFAKFSAYIKNSRPEANEALERGLLKTLQKLDEYLNSPLP
                     DEIDENSMEDIKFSTRKFLDGNEMTLADCNLLPKLHIVKVVAKKYRNFDIPKEMTGIW
                     RYLTNAYSRDEFTNTCPSDKEVEIAYSDVAKRLTK"
ORIGIN

  Query Match              5.5%;  Score 48;  DB 9;  Length 877;
  Best Local Similarity   50.4%;  Pred. No. 0.13;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
Db        531 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 590

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
Db        591 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 650

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
Db        651 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 710
```

```
Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||          ||||  ||   |   ||     |||| ||  ||
Db        711 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 762


RESULT 12
BC012444
LOCUS       BC012444                944 bp    mRNA    linear   PRI 29-JUN-2004
DEFINITION  Homo sapiens chloride intracellular channel 4, mRNA (cDNA clone
            MGC:8812 IMAGE:3861372), complete cds.
ACCESSION   BC012444
VERSION     BC012444.1  GI:15214635
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 944)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-AUG-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: amg@bcm.tmc.edu
            Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
            Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
            A.N., Gibbs, R.A.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 21 Row: f Column: 8
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 7330334.
FEATURES             Location/Qualifiers
     source          1. .944
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:8812 IMAGE:3861372"
                     /tissue_type="Ovary, adenocarcinoma"
                     /clone_lib="NIH_MGC_66"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .944
                     /gene="CLIC4"
                     /note="synonyms: DKFZP566G223, H1, p64H1, huH1, CLIC4L"
                     /db_xref="LocusID:25932"
                     /db_xref="MIM:606536"
     CDS             151. .912
                     /gene="CLIC4"
                     /codon_start=1
                     /product="chloride intracellular channel 4"
                     /protein_id="AAH12444.1"
                     /db_xref="GI:15214636"
                     /db_xref="LocusID:25932"
                     /db_xref="MIM:606536"
                     /translation="MALSMPLNGLKEEDKEPLIELFVKAGSDGESIGNCPFSQRLFMI
                     LWLKGVVFSVTTVDLKRKPADLQNLAPGTHPPFITFNSEVKTDVNKIEEFLEEVLCPP
                     KYLKLSPKHPESNTAGMDIFAKFSAYIKNSRPEANEALERGLLKTLQKLDEYLNSPLP
                     DEIDENSMEDIKFSTRKFLDGNEMTLADCNLLPKLHIVKVVAKKYRNFDIPKEMTGIW
                     RYLTNAYSRDEFTNTCPSDKEVEIAYSDVAKRLTK"
ORIGIN

  Query Match             5.5%;  Score 48;  DB 9;  Length 944;
  Best Local Similarity   50.4%;  Pred. No. 0.13;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        681 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 740

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |    | | |      | ||     || |||||||||  |  |  | |||    |
Db        741 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 800

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||  ||  |    | ||||||| |||  |  |  || |   | |  ||||||  ||
Db        801 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 860

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||          ||||  ||   |   ||     |||| ||  ||
Db        861 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 912


RESULT 13
AF097330
LOCUS       AF097330                999 bp    mRNA    linear   PRI 14-JUN-1999
DEFINITION  Homo sapiens H1 chloride channel mRNA, complete cds.
ACCESSION   AF097330
VERSION     AF097330.1  GI:5052201
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 999)
  AUTHORS   Edwards,J.C.
  TITLE     A novel p64-related Cl- channel: subcellular distribution and
            nephron segment-specific expression
  JOURNAL   Am. J. Physiol. 276 (3 Pt 2), F398-F408 (1999)
  MEDLINE   99170604
   PUBMED   10070163
REFERENCE   2  (bases 1 to 999)
  AUTHORS   Edwards,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-OCT-1998) Medicine, Washington University, 216 S.
            Kingshighway, St. Louis, MO 63110, USA
FEATURES             Location/Qualifiers
     source          1. .999
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
```

```
                     /db_xref="taxon:9606"
                     /cell_line="PancI"
     CDS             141. .902
                     /note="chloride channel of intracellular membranes; p64H1;
                     CLIC4"
                     /codon_start=1
                     /product="H1 chloride channel"
                     /protein_id="AAD38446.1"
                     /db_xref="GI:5052202"
                     /translation="MALSMPLNGLKEEDKEPLIELFVKAGSDGESIGNCPFSQRLFMI
                     LWLKGVVFSVTTVDLKRKPADLQNLAPGTHPPFITFNSEVKTDVNKIEEFLEEVLCPP
                     KYLKLSPKHPESNTAGMDIFAKFSAYIKNSSAEANEALERGLLKTLQKLDEYLNSPLP
                     DEIDENSMEDIKFSTRKFLDGNEMTLADCNLLPKLHIVKVVAKKYRNFDIPKEMTGIW
                     RYLTNAYSRDEFTNTCPSDKEVEIAYSDVAKRLTK"
ORIGIN

  Query Match             5.5%;  Score 48;  DB 9;  Length 999;
  Best Local Similarity   50.4%;  Pred. No. 0.13;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        671 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 730

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |    | | |      |  ||     || |||||||| |  |  | |||    |
Db        731 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 790

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  |  || |   | |  ||||||   ||
Db        791 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 850

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||         ||||  ||   |   ||     ||||  ||  ||
Db        851 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 902


RESULT 14
AX833868
LOCUS       AX833868               2238 bp    DNA     linear   PAT 15-DEC-2003
DEFINITION  Sequence 992 from Patent EP1347046.
ACCESSION   AX833868
VERSION     AX833868.1  GI:39920003
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
            Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
            Masuho,Y.
  TITLE     Full-length cDNA sequences
  JOURNAL   Patent: EP 1347046-A 992 24-SEP-2003;
            Research Association for Biotechnology (JP)
FEATURES             Location/Qualifiers
     source          1. .2238
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             5.5%;  Score 48;  DB 6;  Length 2238;
  Best Local Similarity   50.4%;  Pred. No. 0.12;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        637 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 696

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |    | | |      |  ||     || |||||||| |  |  | |||    |
Db        697 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 756

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  |  || |   | |  ||||||   ||
Db        757 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 816

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||         ||||  ||   |   ||     ||||  ||  ||
Db        817 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 868


RESULT 15
AK095959
LOCUS       AK095959               2238 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ38640 fis, clone HHDPC2003472, highly similar
            to CHLORIDE INTRACELLULAR CHANNEL PROTEIN 4.
ACCESSION   AK095959
VERSION     AK095959.1  GI:21755327
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
            Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
  TITLE     Complete sequencing and characterization of 21,243 full-length
            human cDNAs
  JOURNAL   Nat. Genet. 36 (1), 40-45 (2004)
   PUBMED   14702039
REFERENCE   2
  AUTHORS   Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H.,
            Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
            Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
            Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
            Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2238)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
```

```
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .2238
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HHDPC2003472"
                     /cell_type="dermal papilla cells (HDPC)"
                     /clone_lib="HHDPC2"
                     /note="cloning vector: pME18SFL3-primary culture, dermal
                     papilla cells"
ORIGIN

  Query Match             5.5%;  Score 48;  DB 9;  Length 2238;
  Best Local Similarity   50.4%;  Pred. No. 0.12;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        637 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 696

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     |  | |       |  ||     || |||||||| |  |  | |||    |
Db        697 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 756

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |   | ||||||| |||  |  |  || |   | |  ||||||   ||
Db        757 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 816

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||         ||||  ||   |   ||     |||| ||  ||
Db        817 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 868


Search completed: March 10, 2005, 01:21:16
Job time : 4126 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 23:50:09 ; Search time 3557 Seconds
(without alignments)
9342.172 Million cell updates/sec

Title: US-10-612-379-1
Perfect score: 873
Sequence: 1 atggcagaagcttaccagat.....ccgatgttaatgttcattaa 873

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	312	35.7	701	5	BQ548091	BQ548091 rd20a06.y
2	292.8	33.5	776	6	CA940121	CA940121 ru78d02.y
3	283.8	32.5	759	2	BF718901	BF718901 SWOv3MCAM
4	234.4	26.8	556	6	CD419508	CD419508 rj11f07.y
5	229.4	26.3	599	6	CB375433	CB375433 rw06g06.y
6	209.4	24.0	671	2	BF400411	BF400411 SWOv3MCAM
7	203.6	23.3	567	2	BE029008	BE029008 kp23f02.y
8	202	23.1	483	1	AI043403	AI043403 BSBmMFSZ0
9	189	21.6	566	7	CK242284	CK242284 rx08a04.y
10	185.8	21.3	480	6	CB374401	CB374401 ru99h02.y
11	176.4	20.2	396	4	BI741942	BI741942 kt83b07.y
12	170.2	19.5	450	7	CK349229	CK349229 hggfha10D
13	152.6	17.5	567	6	CA868923	CA868923 pw91g06.y
14	146.8	16.8	426	4	BG225023	BG225023 kp61c08.y
15	138.6	15.9	324	7	CV508408	CV508408 kc71g05.y
16	136.6	15.6	327	1	AA406653	MBAFCZ4D1
17	136.2	15.6	347	6	CB375038	CB375038 ru96d04.y
18	134	15.3	524	5	BM900397	BM900397 rc25b05.y
19	133.6	15.3	400	4	BI749342	BI749342 ro77a04.y
20	130.8	15.0	561	7	CV127561	CV127561 CSEQSDA01
21	126.8	14.5	451	4	BI073297	BI073297 kt29c05.y
22	118.6	13.6	434	2	BE496719	BE496719 Hc_d11_08
23	117.8	13.5	359	6	CB098740	CB098740 ks12d02.y
24	112.6	12.9	468	5	BQ693022	BQ693022 pt10a09.y
25	89	10.2	634	7	CF530481	CF530481 AEL460 ad
26	81.2	9.3	789	7	CN764165	CN764165 ID0AAA9AF
27	80.4	9.2	457	4	BM643363	BM643363 170006873
28	78.8	9.0	630	4	BM609888	BM609888 170006871
29	78.8	9.0	685	7	CR533981	CR533981 CR533981
30	76.4	8.8	403	7	CO273757	CO273757 EK112114.
31	74.4	8.5	572	8	BH761266	BH761266 BMBAC320E
32	72.2	8.3	700	7	CV510881	CV510881 kc64d12.y
33	71.2	8.2	444	7	CO156493	CO156493 EN08162.5
34	70.8	8.1	604	8	BH761283	BH761283 BMBAC320F
35	70.6	8.1	924	6	CD779293	CD779293 EST650654
36	69.8	8.0	909	6	CD794560	CD794560 EST665921
37	67.4	7.7	277	6	CA868922	CA868922 pw91g05.y
38	66.6	7.6	706	4	BM587858	BM587858 170006873
39	66.6	7.6	754	5	BX623011	BX623011 BX623011
40	65	7.4	739	5	BX625406	BX625406 BX625406
41	63.6	7.3	828	6	CD791336	CD791336 EST662697
42	62.8	7.2	598	7	CO279118	CO279118 EK133555.
c 43	61.8	7.1	527	2	AW943833	AW943833 LD46306.3
44	61	7.0	756	5	BP121577	BP121577 BP121577
45	59	6.8	1095	1	AU176220	AU176220 AU176220

ALIGNMENTS

RESULT 1
BQ548091
LOCUS BQ548091 701 bp mRNA linear EST 17-JUN-2002
DEFINITION rd20a06.y1 Meloidogyne incognita egg SL1 TOPO v1 Meloidogyne incognita cDNA 5' similar to WP:CE25703 Y105E8C.A ;contains Alu repetitive element;, mRNA sequence.
ACCESSION BQ548091
VERSION BQ548091.1 GI:21432594
KEYWORDS EST.
SOURCE Meloidogyne incognita (southern root-knot nematode)
ORGANISM Meloidogyne incognita
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE 1 (bases 1 to 701)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Meloidogyne incognita eggs were provided by Andrew Kloek of Divergence Inc., St. Louis, MO.
Seq primer: -40RP from Gibco
High quality sequence stop: 534.
FEATURES Location/Qualifiers
source 1. .701
/organism="Meloidogyne incognita"
/mol_type="mRNA"
/db_xref="taxon:6306"
/dev_stage="egg"
/lab_host="DH10B (Invitrogen)"
/clone_lib="Meloidogyne incognita egg SL1 TOPO v1"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI; Site_2: EcoRI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. cDNA PCR

products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. Meloidogyne incognita eggs were provided by Andrew Kloek of Divergence Inc., St. Louis, MO."

ORIGIN

Query Match 35.7%; Score 312; DB 5; Length 701;
Best Local Similarity 68.7%; Pred. No. 7.5e-80;
Matches 445; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy         44 CAAAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAG 103
Db         54 CAAAACCTATGTTGGAGCTGTATGTTAAGGCATCAGGTATTGATAGCCGAAGAATTGGCG 113

Qy        104 CCGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTG 163
Db        114 CGTGTCTCTTTTGCCAAGAATTTTGGATGGAATTGTATGCTCTCTATGAGGTTGGATGTG 173

Qy        164 CACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCG 223
Db        174 TTCGTGTGGAGGTAAAGACTGTAAATGTCAATTCAGAAGGATTTAAACGCAACTTTTCTG 233

Qy        224 GAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAG 283
Db        234 GAGCTCAACCTCCTATTATGGTGGAAGAAGAAAAAGGGACAACTTATGCTGATAACAGTG 293

Qy        284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
Db        294 AAATTGAACGTCGAATCTTCCATCTCTGTCAAGGTTTTAATGTTCCGCTTTTTGAGAAGG 353

Qy        344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
Db        354 ATTTACAAGTTGAGAAAACTATCCAGAATCTTTATCGAAATTTCAAAATCTTTCTGCAAG 413

Qy        404 CAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCC---ATCGAGAGTTGAAGATCTTCCAG 460
Db        414 CAAAGGCAAATCATGACAAAGGTAAAAGACATCCTAGTGCAATAGAAGAGGGCCTTCCCC 473

Qy        461 CACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGC 520
Db        474 CACAGGTGACATCTTCACACAATAAATTATTGGAGCAATTGGCCAATATTGACCAGTTGT 533

Qy        521 TATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAAC 580
Db        534 TGGCTGATCGTTCTTCTCGTTATCTTCTAAGTGAATCAATGACTGAATATGACTGTGAGA 593

Qy        581 TGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTC 640
Db        594 TTATGCCACGTTTGCATCATATTCGTATCATTGGAGAACGCCTTTTGAATTTTTACATTC 653

Qy        641 CACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTA 688
Db        654 CACGTCAATTCACTCATTTATGGGCATATATATTAACTGCATATCGGA 701
```

RESULT 2
CA940121
LOCUS CA940121 776 bp mRNA linear EST 30-DEC-2002
DEFINITION ru78d02.y1 Heterodera glycines virgin female Heterodera glycines cDNA 5' similar to WP:CE25711 Y105E8E.Y ;, mRNA sequence.
ACCESSION CA940121
VERSION CA940121.1 GI:27428601
KEYWORDS EST.
SOURCE Heterodera glycines
ORGANISM Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
REFERENCE 1 (bases 1 to 776)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This library was generated by cloning cDNAs directionally into Uni-ZAP(Stratagene) (T3 primer/EcoRI are at the 5'-end and T7/XhoI are at the 3'-end). The library was excised [now in pBluescript SK(+)] and normalized (Bonaldo et al 1996 Genome Research 6:791-806). Library constructed by Thomas Baum (tbaum@iastate.edu), Iowa State University, Plant Pathology Department and Jeff McDermott (jpmcderm@iastate.edu).
Seq primer: T3 from Gibco
High quality sequence stop: 488.
FEATURES Location/Qualifiers
source 1. .776
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/sex="female"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Heterodera glycines virgin female"
/note="Vector: pBluescript SK+ (Stratagene); Site_1: XhoI; Site_2: EcoRI; This library was generated by cloning cDNAs directionally into Uni-ZAP(Stratagene) (T3 primer/EcoRI are at the 5'-end and T7/XhoI are at the 3'-end). The library was excised [now in pBluescript SK(+)] and normalized (Bonaldo et al 1996 Genome Research 6:791-806). Library constructed by Thomas Baum (tbaum@iastate.edu), Iowa State University, Plant Pathology Department and Jeff McDermott (jpmcderm@iastate.edu)."

ORIGIN

Query Match 33.5%; Score 292.8; DB 6; Length 776;
Best Local Similarity 63.6%; Pred. No. 3.3e-74;
Matches 463; Conservative 0; Mismatches 262; Indels 3; Gaps 1;

```
Qy         44 CAAAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAG 103
Db         48 CAAAGCCTTTGCTCGAGTTGTATGTGAAGGCCTCAGGGATTGACAGCCGACGCATCGGTG 107

Qy        104 CCGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTG 163
Db        108 CATGCTTATTTTGCCAAGAATTTTGGATGGAACTCTACGCTCTTTACGAAGTGGGCTGTG 167

Qy        164 CACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCG 223
Db        168 TTCGTGTTGAGGTAAAAACTGTGAATGTCAATTCGGAAGGATACAAAAAGCACTTTTCGG 227

Qy        224 GAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAG 283
Db        228 GAGCACAGCCGCCAATAATGGTGGAACAAGACAAAGGCACGGCAAATGCGGACAACACTG 287

Qy        284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
Db        288 ACATCGAGCGACGCATTTTCCAACTTTGCAACGAATTCCATGTGCCATTGTTCGAGAAGG 347

Qy        344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
Db        348 ATCCGGAAGTGGCAAAGACCATCCAGGATTTGTACAGGAACTTCAAAATATTTCTCCAAG 407

Qy        404 CAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCAC 463
Db        408 CCAAAACGAATTTGGACAAAGGCAAGCAGCAACCCAGCTCCATTGAAGGACTGCCGCCCC 467
```

```
Qy        464 AGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTAT 523
              |  |           |||    | |    || || ||  | |||||  |||||||| |
Db        468 AAGTCGCCTCATCGCACAGCAAATTGGTAGAACAGCTGGCGAATATAAATCAGTTGTTGG 527

Qy        524 CCGAG---AGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAAC 580
              |||||    | ||  | || || || ||  | |  |||| ||||| ||||| ||
Db        528 CCGAGCGTGGCAACACACGCTACCTTCTCGGCGAGACGATGACGGAATACGACTGCGAGT 587

Qy        581 TGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTC 640
              |||||||||| ||||||||  | ||  |  | ||   |  ||| || || |||   ||||
Db        588 TGATGCCACGCCTTCATCACCTGCGCGTGGTCGGGGAGCGACTGCTCGGCTTCACCATTC 647

Qy        641 CACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTA 700
              | ||  |  ||||||| || |||   ||| | || ||  | ||| | || || || || |
Db        648 CGCACCAACTCACTCACCTGTGGAACTATGTGCTGACCTCGTACAGAACCGCGGCCTTCA 707

Qy        701 TTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGT 760
              | |||   |||||||| || |||||||| | || || || || || ||| | || ||
Db        708 TCGAGTCGTGTCCCGCTGACCAGGACATTCTCCACCATTACAAGGAGCAACTAAACCTTA 767

Qy        761 TCACAAAT 768
              ||   ||
Db        768 GCATCGAT 775


RESULT 3
BF718901
LOCUS       BF718901                759 bp    mRNA    linear   EST 03-JAN-2001
DEFINITION  SWOv3MCAM58D06SK Onchocerca volvulus molting L3 larva cDNA
            (SL96MLW-OvmL3) Onchocerca volvulus cDNA clone SWOv3MCAM58D06 5',
            mRNA sequence.
ACCESSION   BF718901
VERSION     BF718901.1  GI:12019909
KEYWORDS    EST.
SOURCE      Onchocerca volvulus
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
  TITLE     Genes expressed in molting L3 larvae of Onchocerca volvulus
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Onchocerca volvulus"
                     /mol_type="mRNA"
                     /strain="Kumba, Cameroons"
                     /db_xref="taxon:6282"
                     /clone="SWOv3MCAM58D06"
                     /dev_stage="molting L3"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="Onchocerca volvulus molting L3 larva cDNA
                     (SL96MLW-OvmL3)"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. Third-stage
                     larvae, L3,  were isolated from infected black flies in
                     Cameroon (forest strain). The L3 were cultured in 20% FCS
                     in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
                     culture. L3 of O. volvulus molt to fourth-stage larvae by
                     day 5 in culture. mRNA was isolated from approximately
                     6000 molting larvae (mL3),2000 larvae from day 1, 2 or 3
                     in culture, and converted to double-stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNA pol I. The library was constructed in the lambda
                     Uni-Zap XR vector and has 1 x 10E6 independent
                     recombinants and the average insert size is ~1200 bp. The
                     library was constructed by Sara Lustigman and Michelle
                     Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.
                     The library is available from Dr. Sara Lustigman (email:
                     slustigm@nybc.org)."
ORIGIN

  Query Match           32.5%;  Score 283.8;  DB 2;  Length 759;
  Best Local Similarity   68.9%;  Pred. No. 1.4e-71;
  Matches  450;  Conservative    0;  Mismatches  192;  Indels   11;  Gaps    4;

Qy         46 AAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCC 105
              ||||| ||||| ||||| || || || || || ||||| ||   ||| ||||||| ||
Db         97 AAACCACTTCTGGAGCTTTATGTCAAGGCTTCGGGAATCGAAAATCGGCGCATTGGTGCA 156

Qy        106 GATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCA 165
                 | || ||||| ||||| |||||||| | || ||||||||||||||||| ||||
Db        157 TGCTTGTTTTGTCAAGAATTTTGGATGGAACTTTACGCTCTTTATGAGATTGGCGTTGTT 216

Qy        166 CGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGA 225
              ||||| |||||||||| || ||| | |||||||||||||| || ||||  |||||||||
Db        217 CGAGTTGAAGTGAAGACAGTAAACATCAATTCTGAAGCATTCAAAAAGAGTTTTCTCGGA 276

Qy        226 GCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAG 285
              || ||||| || |||||| | ||| |  |||| |     || || ||||||||||| ||
Db        277 GCTCAACCGCCAATTATGGTGGAAAATAAAAATG---CAACTTATACTGATAATCGTGAA 333

Qy        286 ATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGAT 345
              || |||||||| || |||||||| || || || |||||||| ||| | |||| || |||
Db        334 ATCGAAGGACGCATTTTTCATTTAGCGAAAGAGTTCAATGTGCCATTGTTTGAGAAAGAT 393

Qy        346 CCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCA 405
              |||    |  || ||  |||| || |  |||||||| || ||||||  | ||  ||||| |
Db        394 CCAGTGGTAGAAAAACGAATTGAAAGTTTGTACAGAAATTTCAAAATCTTTTTGCGATCC 453

Qy        406 AAAGTAGAGTTCGATAAGGGAAAAAAGG------AGCCATCGAGAGTTGAAGATCTTCCA 459
              |||   |||   || | || |  ||| |        ||| |   | | ||    ||||||
Db        454 AAAAGTGAGCATGACAGGGAACGAAAAGGATCATCACCAACACCAATGGATTCGCTTCCA 513

Qy        460 GCACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTG 519
              | |||   ||||    || ||   | |   ||| || || || || ||||||||  |
Db        514 CCGCAGCAAAAAGCATCATATAACAAACTTGTTGAACAGCTGTCAAACATTGATCAACTT 573

Qy        520 CTATCCGAGA-GAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGA 578
              ||    |||  |||| || |||||||| || ||| | |          || ||||| || ||
Db        574 CTGAGTGAGCCGAAACTCACGATATCTTCTGGGACAAAAGTATGACGAGTATGATTGCGA 633

Qy        579 ACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATAT 638
              ||| ||||| |||||||||||| |||||||  |||||       || || || || || ||
Db        634 ACTTATGCCTCGTCTTCATCACATTCGAATCGTTGGA-CACGTCTGCTCGGTTTTGACAT 692

Qy        639 TCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAG 691
              ||| |  ||| |  ||  ||| |||   ||| |  | | ||||| ||||||  |
Db        693 TCCGCTGAATCTACCTATTCTTTGGAACTATGTGTTGAATGCATCCCGTACCG 745


RESULT 4
CD419508
LOCUS       CD419508                556 bp    mRNA    linear   EST 02-JUN-2003
DEFINITION  rj11f07.y1 Meloidogyne chitwoodi egg pAMP1 v1 Meloidogyne chitwoodi
            cDNA 5' similar to WP:CE25711 Y105E8E.Y ;, mRNA sequence.
ACCESSION   CD419508
VERSION     CD419508.1  GI:31325771
KEYWORDS    EST.
SOURCE      Meloidogyne chitwoodi
  ORGANISM  Meloidogyne chitwoodi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
```

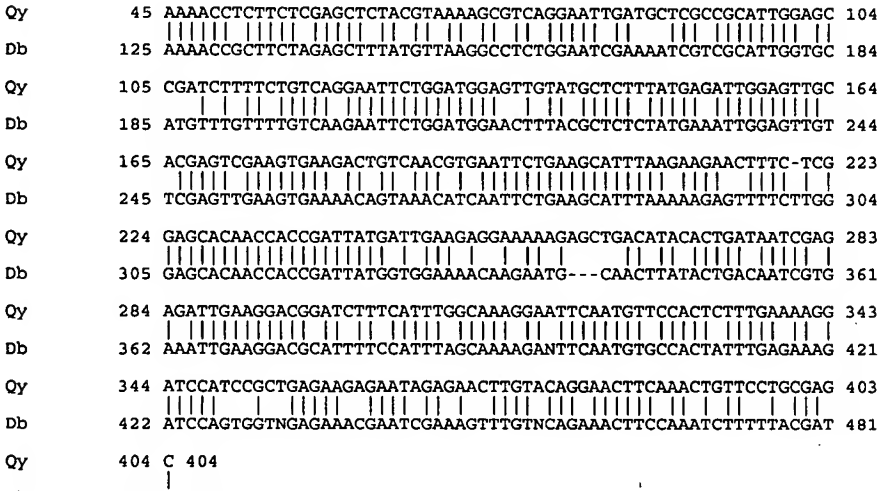
```
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. Makedonka
            Dautova at Washington University, St. Louis. The cDNA was made by
            using Dynabead oligo-dT priming (Dynal). PCR based library using a
            modified protocol from the SMART PCR cDNA Synthesis Kit from
            Clontech. Directionally cloned into the UDG sites of pAMP1.
            Meloidogyne chitwoodi eggs were provided by Dr. David Bird at North
            Carolina State University, Raleigh, NC (david_bird@ncsu.edu).
            Seq primer: T3 from Gibco.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Meloidogyne chitwoodi"
                     /mol_type="mRNA"
                     /db_xref="taxon:59747"
                     /dev_stage="egg"
                     /lab_host="DH10B"
                     /clone_lib="Meloidogyne chitwoodi egg pAMP1 v1"
                     /note="Vector: pAMP1; Site_1: Not I; Site_2: Sal I; The
                     library was constructed by Claire Murphy and Dr. Makedonka
                     Dautova at Washington University, St. Louis. The cDNA was
                     made by using Dynabead oligo-dT priming (Dynal). PCR based
                     library using a modified protocol from the SMART PCR cDNA
                     Synthesis Kit from Clontech. Directionally cloned into the
                     UDG sites of pAMP1. Meloidogyne chitwoodi eggs were
                     provided by Dr. David Bird at North Carolina State
                     University, Raleigh, NC (david_bird@ncsu.edu)."
ORIGIN

  Query Match             26.8%;  Score 234.4;  DB 6;  Length 556;
  Best Local Similarity   65.1%;  Pred. No. 3.7e-57;
  Matches  362;  Conservative    0;  Mismatches  191;  Indels    3;  Gaps    1;

Qy        207 TAAGAAGAACTTTCTCGGAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGAC 266
              ||||   || |||   || || ||||| |||||||| | || || ||||| |     ||
Db          1 TAAGCGCAATTTTTCTGGTGCTCAACCTCCGATTATGGTCGAGGAAGAAAAGGCAACAAC 60

Qy        267 ATACACTGATAATCGAGAGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGT 326
              ||  | || ||  | || |||||| | || || || ||| |     | |  || |||||
Db         61 CTATGCAGACAACAGCGAAATTGAACGTCGTATTTTCCATCTCTGTCAAGGTTTTAATGT 120

Qy        327 TCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTT 386
              |||||| ||||| ||||||       | |||||| |  ||  | ||  | ||    || ||
Db        121 TCCACTTTTTGAGAAGGATTTGCAAGTTGAGAAAACTATTCAAAATCTTTATCTTAATTT 180

Qy        387 CAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGT 446
              |||| | ||  ||| |||||||         || || || |||    ||||    | | |
Db        181 CAAAATCTTTTTGCAAGCAAAAGCGACTCATGACAAAGGCAAACGACAGCCTAGTACAAT 240

Qy        447 TGAAGA---TCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATC 503
              || |||   |||||| ||||| |   |  |    | |||  | | |  |||||| |   |
Db        241 TGGAGAGGGTCTTCCACCACAGGTGGCAACTTCTCATAATAAATTATTGGAGCAATTGGC 300

Qy        504 CAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGAC 563
                | |||| |||||  |  | ||  |    ||||| ||||| ||  |  |    |||||
Db        301 TGGTGTTGACCAGTTATTGGCTGATCGTGCTTCTCGTTATCTTCTAAGCGAATCAATGAC 360

Qy        564 TGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACT 623
              |||||||| |||||| | |||||||| | |||||||| || |||||||| |   ||
Db        361 TGAATATGATTGTGAAGTTATGCCACGTTTGCATCATATACGTATTATTGGAGAGCGTCT 420

Qy        624 TCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATA 683
              |||    ||  | |||||||   | || || ||| | ||||| |||||  | ||||||||
Db        421 TCTCAATTTTTACATTCCACGACAATTTACACATTTATGGGCATATATATTAACTGCATA 480

Qy        684 CCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAA 743
              || || || || ||||||||   |||||| || ||||| || ||| ||||||| |||||
Db        481 TCGAACGGCTGCCTTTATTGAATCTTGTCCAGCTGATCAAGATATTCTTCATCATTATAA 540

Qy        744 AGAACAAATGAATCTG 759
              || ||| | ||| ||
Db        541 GGAGCAATTAAATTTG 556


RESULT 5
CB375433
LOCUS       CB375433                599 bp    mRNA    linear   EST 18-MAR-2003
DEFINITION  rw06g06.y1 Heterodera glycines virgin female Heterodera glycines
            cDNA 5' similar to WP:CE25703 Y105E8C.A ;, mRNA sequence.
ACCESSION   CB375433
VERSION     CB375433.1  GI:29051812
KEYWORDS    EST.
SOURCE      Heterodera glycines
  ORGANISM  Heterodera glycines
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This library was generated by cloning cDNAs directionally into
            Uni-ZAP(Stratagene) (T3 primer/EcoRI are at the 5'-end and T7/XhoI
            are at the 3'-end). The library was excised [now in pBluescript
            SK(+)] and normalized (Bonaldo et al 1996 Genome Research
            6:791-806). Library constructed by Thomas Baum (tbaum@iastate.edu),
            Iowa State University, Plant Pathology Department and Jeff
            McDermott (jpmcderm@iastate.edu).
            Seq primer: -40RP from Gibco
            High quality sequence stop: 450.
FEATURES             Location/Qualifiers
     source          1. .599
                     /organism="Heterodera glycines"
                     /mol_type="mRNA"
                     /db_xref="taxon:51029"
                     /sex="female"
                     /tissue_type="whole organism"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Heterodera glycines virgin female"
                     /note="Vector: pBluescript SK+ (Stratagene); Site_1: XhoI;
                     Site_2: EcoRI; This library was generated by cloning cDNAs
                     directionally into Uni-ZAP(Stratagene) (T3 primer/EcoRI
                     are at the 5'-end and T7/XhoI are at the 3'-end). The
                     library was excised [now in pBluescript SK(+)] and
                     normalized (Bonaldo et al 1996 Genome Research 6:791-806).
                     Library constructed by Thomas Baum (tbaum@iastate.edu),
                     Iowa State University, Plant Pathology Department and Jeff
                     McDermott (jpmcderm@iastate.edu)."
ORIGIN

  Query Match             26.3%;  Score 229.4;  DB 6;  Length 599;
```

```
  Best Local Similarity   62.9%;  Pred. No. 1.1e-55;
  Matches  372;  Conservative    0;  Mismatches  216;  Indels    3;  Gaps    1;

Qy          44 CAAAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAG 103
               |||| ||| | |||||| | || || || || || ||||| |||||    || ||||| || |
Db           9 CAAAGCCTTTGCTCGAGTTGTATGTGAAGGCCTCAGGGATTGACAGCCGACGCATCGGTG 68

Qy         104 CCGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTG 163
               |      | || || || ||||| |||||||| | || |||||||| ||  | ||    ||
Db          69 CATGCTTATTTTGCCAAGAATTTTGGATGGAACTCTACGCTCTTTACGAAGTGGGCTGTG 128

Qy         164 CACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCG 223
                 || || || || ||||| || || ||||| |||| ||  || ||| |||||    |
Db         129 TTCGTGTTGAGGTAAAAACTGTGAATGTCAATTCGGAAGGATACAAAAAGCACTTTTCGG 188

Qy         224 GAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAG 283
               ||||||| || || ||| | |||    || ||||    | || |  | || ||    |
Db         189 GAGCACAGCCGCCAATAATGGTGGAACTTGACAAAGGCACGGCAAATGCGGACAACACTG 248

Qy         284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
               | || ||  |||| || || ||  |    || |||||| |||| ||| | || || ||||
Db         249 ACATCGAGCGACGCATTTTCCAACTTTGCAACGAATTCCATGTGCCATTGTTCGAGAAGG 308

Qy         344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
               ||||    |  |  |||| ||  || | ||||||||||||||||| | || || | ||
Db         309 ATCCGGAAGTGGCAAAGACCATCCAGGATTTGTACAGGAACTTCAAAATATTTCTCCAAG 368

Qy         404 CAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCAC 463
               | |||    | || || || || ||  || | ||        ||||||  || ||  | |
Db         369 CCAAAACGAATTTGGACAAAGGCAAGCAGCAACCCAGCTCCATTGAAGGACTGCCGCCCC 428

Qy         464 AGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTAT 523
               |  |            |||    | |    || || ||  | |||||  |||||||| |
Db         429 AAGTCGCCTCATCGCACAGCAAATTGGTAGAACAGCTGGCGAATATAAATCAGTTGTTGG 488

Qy         524 CCGAG---AGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAAC 580
               |||||    | ||  | || || || || ||  | |  |||||  |||||  ||||| ||
Db         489 CCGAGCGTGGCAACACACGCTACCTTCTCGGCGAGACGATGACGGAATACGACTGCGAGT 548

Qy         581 TGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGAT 631
               ||||||||||  |||||||  | ||  |  | ||   |  ||| || || |
Db         549 TGATGCCACGCCTTCATCACCTGCGCGTGGTCGGGGAGCGACTGCTCGGCT 599


RESULT 6
BF400411
LOCUS       BF400411                671 bp    mRNA    linear   EST 28-NOV-2000
DEFINITION  SWOv3MCAM45F09SK Onchocerca volvulus molting L3 larva cDNA
            (SL96MLW-OvmL3) Onchocerca volvulus cDNA clone SWOv3MCAM45F09 5',
            mRNA sequence.
ACCESSION   BF400411
VERSION     BF400411.1  GI:11386455
KEYWORDS    EST.
SOURCE      Onchocerca volvulus
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 671)
  AUTHORS   Williams,S.A., Lizotte-Waniewski,M., Laney,S. and Lustigman,S.
  TITLE     Genes expressed in molting L3 larvae of Onchocerca volvulus
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .671
                     /organism="Onchocerca volvulus"
                     /mol_type="mRNA"
                     /strain="Kumba, Cameroons"
                     /db_xref="taxon:6282"
                     /clone="SWOv3MCAM45F09"
                     /dev_stage="molting L3"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="Onchocerca volvulus molting L3 larva cDNA
                     (SL96MLW-OvmL3)"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. Third-stage
                     larvae, L3,  were isolated from infected black flies in
                     Cameroon (forest strain). The L3 were cultured in 20% FCS
                     in IMDM+ NCTC 135 and collected after day 1, 2, or 3 in
                     culture. L3 of O. volvulus molt to fourth-stage larvae by
                     day 5 in culture. mRNA was isolated from approximately
                     6000 molting larvae (mL3),2000 larvae from day 1, 2 or 3
                     in culture, and converted to double-stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNA pol I. The library was constructed in the lambda
                     Uni-Zap XR vector and has 1 x 10E6 independent
                     recombinants and the average insert size is ~1200 bp. The
                     library was constructed by Sara Lustigman and Michelle
                     Lizotte-Waniewski in the Laboratory of Dr. S. A. Williams.
                     The library is available from Dr. Sara Lustigman (email:
                     slustigm@nybc.org)."
ORIGIN

  Query Match           24.0%;  Score 209.4;  DB 2;  Length 671;
  Best Local Similarity   75.5%;  Pred. No. 8.3e-50;
  Matches  274;  Conservative    0;  Mismatches   86;  Indels    3;  Gaps    1;

Qy          46 AAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGCC 105
               ||||| ||||| ||||| || || || || || ||||| ||   ||| |||||||| ||
Db          97 AAACCACTTCTGGAGCTTTATGTCAAGGCTTCGGGAATCGAAAATCGGCGCATTGGTGCA 156

Qy         106 GATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGCA 165
                  | || ||||| ||||| ||||||||  | || ||||||||||||||||| ||||
Db         157 TGCTTGTTTTGTCAAGAATTTTGGATGGAACTTTACGCTCTTTATGAGATTGGCGTTGTT 216

Qy         166 CGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGA 225
               ||||| |||||||||||| || ||| | |||||||||||||| || ||||  |||||||||
Db         217 CGAGTTGAAGTGAAGACAGTAAACATCAATTCTGAAGCATTCAAAAAGAGTTTTCTCGGA 276

Qy         226 GCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAGAG 285
               || ||||| || |||||| | ||| |  |||| |     || || |||||||||| ||
Db         277 GCTCAACCGCCAATTATGGTGGAAAATAAAAATG---CAACTTATACTGATAATCGTGAA 333

Qy         286 ATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGAT 345
               || |||||||| || |||||||| || || || |||||||| ||| | ||||| || |||
Db         334 ATCGAAGGACGCATTTTTCATTTAGCGAAAGAGTTCAATGTGCCATTGTTTGAGAAAGAT 393

Qy         346 CCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCA 405
               |||    |  || ||  |||| || |  |||||||| || |||||| | ||  |||||  |
Db         394 CCAGTGGTAGAAAAACGAATTGAAAGTTTGTACAGAAATTTCAAAATCTTTTTGCGATCC 453

Qy         406 AAA 408
               |||
Db         454 AAA 456


RESULT 7
BE029008
LOCUS       BE029008                567 bp    mRNA    linear   EST 09-MAY-2001
DEFINITION  kp23f02.y1 TBN95TM-SSFH Strongyloides stercoralis cDNA 5' similar
            to WP:F26H11.5 CE15912 ;, mRNA sequence.
ACCESSION   BE029008
VERSION     BE029008.1  GI:8322469
KEYWORDS    EST.
SOURCE      Strongyloides stercoralis
  ORGANISM  Strongyloides stercoralis
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 567)
```

```
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Dr. Thomas Nutman and colleagues of
            NIAID, NIH (tnutman@nih.gov). DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 418.
FEATURES             Location/Qualifiers
     source          1. .567
                     /organism="Strongyloides stercoralis"
                     /mol_type="mRNA"
                     /strain="Filariform larvae obtained from humans"
                     /db_xref="taxon:6248"
                     /lab_host="XL-1 Blue MRF' (Stratagene)"
                     /clone_lib="TBN95TM-SSFH"
                     /note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1:
                     EcoRI; Site_2: XhoI; mRNA was purified from 4 x 10E5
                     filariform larvae which had been isolated from infected
                     humans.  cDNA was constructed and, using adaptors, was
                     cloned  unidirectionally into the vector from the EcoRI
                     site to the XhoI site.  The  library has an unamplified
                     titer of 1.5 x 10E6 pfu/ml and an amplified,  undiluted
                     titer of 7 x 10E9 pfu/ml.  The average insert size of the
                     unamplifed library is 975 bp (range, 500-1500)."
ORIGIN

  Query Match           23.3%;  Score 203.6;  DB 2;  Length 567;
  Best Local Similarity   61.5%;  Pred. No. 3.9e-48;
  Matches  348;  Conservative    0;  Mismatches  209;  Indels    9;  Gaps    1;

Qy        190 GTGAATTCTGAAGCATTTAAGAAGAACTTTCTCGGAGCACAACCACCGATTATGATTGAA 249
              || |||||||||||| ||||| ||    ||||| ||    ||||| || ||||||||||||
Db          1 GTTAATTCTGAAGCTTTTAAAAAATCTTTTCTTGGCAGTCAACCTCCAATTATGATTGAA 60

Qy        250 GAGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATCTTTCATTTG 309
              |     |||||    |||||  | |||||| |||| ||||||  |  | || ||||||  |
Db         61 GCTTCCAAAGATGCAACATATTCAGATAATAGAGATATTGAAAGTAGAATATTTCATCTT 120

Qy        310 GCAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAG 369
              || || |||||  ||||  | || |||||||| |||    | | ||| ||||  ||  |
Db        121 GCCAAAGAATTTGATGTGGCTCTTTTTGAAAAAGATATTACTGTTGAAAAGATTATTCAA 180

Qy        370 AACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAA 429
                | || || || ||| | || ||   |||||||   ||    |||||  |  ||
Db        181 TCACTTTATAGAAATTTTAAATTATTTCTTAAAGCAAAAACTGATCATGATAAAGTTAAG 240

Qy        430 AAG---------GAGCCATCGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTAC 480
              |||         |    |   |||| | | |||  |    ||  || |
Db        241 AAGATAAATAGAAATAATACTTTACAAGAAGGTTTACCACAATCTGTTGTAGCTTGTCGT 300

Qy        481 AATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGA 540
              |    | |    || ||| || |    |||||||| ||| || |    |||   ||  ||
Db        301 ACAAAATTACTAGAACAATTAGCATCAATTGATCATTTGTTAGCACGAAGAGGTTCAAGA 360

Qy        541 TATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCAT 600
              ||| || |   | | ||| ||  ||| ||||| |||||| |||||||| |  | ||||||
Db        361 TATTTATTATCACAAAGTTTGTGTGAGTATGATTGTGAATTGATGCCAAGATTACATCAT 420

Qy        601 ATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTC 660
```

```
              ||| ||||| ||||      | |  | |  |  || ||||||||||| | ||  || |
Db        421 ATTAGAATTGTTGGTCAAAGATTATTAGATTATGAAATTCCACATAATGTTACCTATGTA 480

Qy        661 TGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGAT 720
              |||  |||||| || ||||| ||  |  | || || ||   |   ||||| || || |||
Db        481 TGGAATTATATGCTTACTGCCTATAGAGCTGCTGCGTTCGATTGAAGTTGCCCAGCTGAT 540

Qy        721 CAGGACATTATTCATCACTATAAAGA 746
              ||  | ||||| || || || | |||
Db        541 CAACATATTATACACCATTACAGAGA 566
```

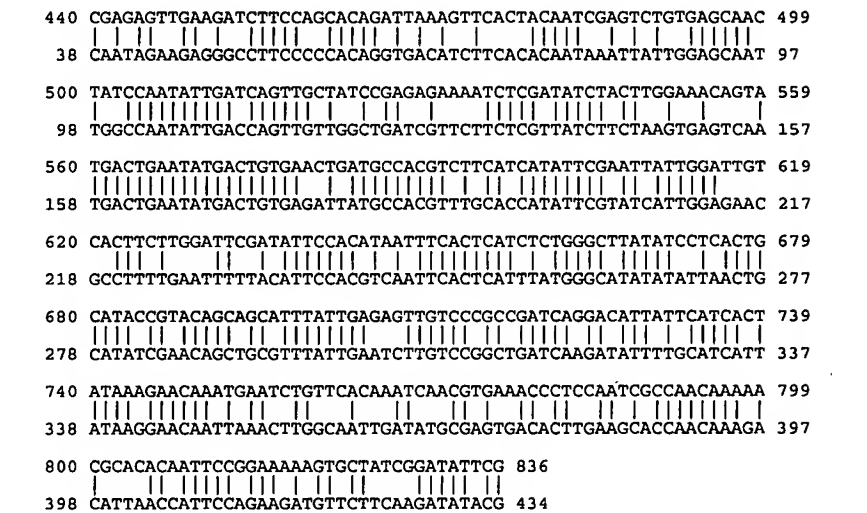
```
RESULT 8
AI043403
LOCUS       AI043403                483 bp    mRNA    linear   EST 01-JUL-1998
DEFINITION  BSBmMFSZ06I10SK Brugia malayi microfilaria cDNA (SAW94LS-BmMf)
            Brugia malayi cDNA clone BSBmMFSZ06I10 5', mRNA sequence.
ACCESSION   AI043403
VERSION     AI043403.1  GI:3286660
KEYWORDS    EST.
SOURCE      Brugia malayi
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in microfilaria of Brugia malayi
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            EST submitted by Molecular Parasitology Group, New England Biolabs,
            Inc., 32 Tozer Road, Beverly, MA. 01915, USA. Email:
            dnaseq@neb.com. SZ in the clone designation, refers to 'Subtracted
            Method Z'. The numerical designation after the SZ refers to the
            microtiter tray number (01 - 48) followed by the letter and number
            of the microtiter tray position. Method 'Z' refers to sequenced
            clones which have been selected after hybridization subtraction
            using highly redundant clones present in the gridded library
            (18,000 mass-excised colonies gridded as a high density array on
            nylon filters). Colonies not represented in the probe sets were
            used as templates for the sequencing reactions.
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Brugia malayi"
                     /mol_type="mRNA"
                     /strain="TRS Labs"
                     /db_xref="taxon:6279"
                     /clone="BSBmMFSZ06I10"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="Brugia malayi microfilaria cDNA
                     (SAW94LS-BmMf)"
                     /note="Vector: lambda UniZap XR; Site_1: EcoR I; Site_2:
                     Xho I; Lymphatic filarial nematode parasite of humans.
                     mRNA was prepared from microfilariae of Brugia malayi
                     isolated from jirds and converted to double stranded cDNA
                     using reverse transcriptase and oligo(dT) followed by
                     RNase H and DNApol I. The library had 3.5 x 10E5
                     independent recombinants and average insert size was 900
                     base pairs. The library was constructed by Lori Saunders.
                     The library is available from Dr. S.A. Williams, email
                     genome@smith.edu."
ORIGIN

  Query Match           23.1%;  Score 202;  DB 1;  Length 483;
  Best Local Similarity   75.9%;  Pred. No. 1.1e-47;
  Matches  274;  Conservative    0;  Mismatches   83;  Indels    4;  Gaps    2;
```



```
Qy         45 AAAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAGC 104
Db        125 AAAACCGCTTCTAGAGCTTTATGTTAAGGCCTCTGGAATCGAAAATCGTCGCATTGGTGC 184

Qy        105 CGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTGC 164
Db        185 ATGTTTGTTTTGTCAAGAATTCTGGATGGAACTTTACGCTCTCTATGAAATTGGAGTTGT 244

Qy        165 ACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTC-TCG 223
Db        245 TCGAGTTGAAGTGAAAACAGTAAACATCAATTCTGAAGCATTTAAAAAGAGTTTTCTTGG 304

Qy        224 GAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAG 283
Db        305 GAGCACAACCACCGATTATGGTGGAAAACAAGAATG---CAACTTATACTGACAATCGTG 361

Qy        284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
Db        362 AAATTGAAGGACGCATTTTCCATTTAGCAAAAGANTTCAATGTGCCACTATTTGAGAAAG 421

Qy        344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
Db        422 ATCCAGTGGTNGAGAAACGAATCGAAAGTTTGTNCAGAAACTTCCAAATCTTTTTACGAT 481

Qy        404 C 404
Db        482 C 482
```

RESULT 9
CK242284
LOCUS CK242284 566 bp mRNA linear EST 11-DEC-2003
DEFINITION rx08a04.y1 Meloidogyne paranaensis egg SMART pGEM Meloidogyne paranaensis cDNA 5' similar to WP:CE25711 Y105E8E.Y ;, mRNA sequence.
ACCESSION CK242284
VERSION CK242284.1 GI:39747874
KEYWORDS EST.
SOURCE Meloidogyne paranaensis
ORGANISM Meloidogyne paranaensis
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE 1 (bases 1 to 566)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Cloned unidirectionally. Poly(A)+ RNA was concentrated and purified using Dynabeads (Dynal) and mRNA eluted for first strand synthesis. First strand cDNA was created using MMLV RT (Powerscript, Clontech) and primed with oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5'SMART 'anchor' added using chimeric DNA-RNA oligo (SMART-NotI-r-GGG). 12 PCR cycles were done using first strand and primers specific to SMART oligo(5'PCR-primer) and 3' end(XhoI-No-dT). Double stranded cDNA was digested using XhoI/NotI, fractioned on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as host cells. Library constructed by Joanna Carlson of North Carolina State University.
Putative full length read
The vector to vector length is 632
Seq primer: Sp6.
FEATURES Location/Qualifiers
source 1. .566
/organism="Meloidogyne paranaensis"
/mol_type="mRNA"
/db_xref="taxon:189293"
/tissue_type="whole organism"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="Meloidogyne paranaensis egg SMART pGEM"
/note="Vector: pGEM-11zf(+) (Promega); Site_1: XhoI; Site_2: NotI; Cloned unidirectionally. Poly(A)+ RNA was concentrated and purified using Dynabeads (Dynal) and mRNA eluted for first strand synthesis. First strand cDNA was created using MMLV RT (Powerscript, Clontech) and primed with oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5'SMART 'anchor' added using chimeric DNA-RNA oligo (SMART-NotI-r-GGG). 12 PCR cycles were done using first strand and primers specific to SMART oligo(5'PCR-primer) and 3' end(XhoI-No-dT). Double stranded cDNA was digested using XhoI/NotI, fractioned on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as host cells. Library constructed by Joanna Carlson."
ORIGIN

Query Match 21.6%; Score 189; DB 7; Length 566;
Best Local Similarity 67.3%; Pred. No. 7.4e-44;
Matches 267; Conservative 0; Mismatches 130; Indels 0; Gaps 0;



```
Qy        440 CGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAAC 499
Db         38 CAATAGAAGAGGGCCTTCCCCCACAGGTGACATCTTCACACAATAAATTATTGGAGCAAT 97

Qy        500 TATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTA 559
Db         98 TGGCCAATATTGACCAGTTGTTGGCTGATCGTTCTTCTCGTTATCTTCTAAGTGAGTCAA 157

Qy        560 TGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGT 619
Db        158 TGACTGAATATGACTGTGAGATTATGCCACGTTTGCACCATATTCGTATCATTGGAGAAC 217

Qy        620 CACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTG 679
Db        218 GCCTTTTGAATTTTTACATTCCACGTCAATTCACTCATTTATGGGCATATATATTAACTG 277

Qy        680 CATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACT 739
Db        278 CATATCGAACAGCTGCGTTTATTGAATCTTGTCCGGCTGATCAAGATATTTTGCATCATT 337

Qy        740 ATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAACAAAAA 799
Db        338 ATAAGGAACAATTAAACTTGGCAATTGATATGCGAGTGACACTTGAAGCACCAACAAAGA 397

Qy        800 CGCACACAATTCCGGAAAAAGTGCTATCGGATATTCG 836
Db        398 CATTAACCATTCCAGAAGATGTTCTTCAAGATATACG 434
```

RESULT 10
CB374401
LOCUS CB374401 480 bp mRNA linear EST 18-MAR-2003
DEFINITION ru99h02.y1 Heterodera glycines virgin female Heterodera glycines cDNA 5' similar to WP:CE25703 Y105E8C.A ;, mRNA sequence.
ACCESSION CB374401
VERSION CB374401.1 GI:29049758
KEYWORDS EST.
SOURCE Heterodera glycines
ORGANISM Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
REFERENCE 1 (bases 1 to 480)

```
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This library was generated by cloning cDNAs directionally into
            Uni-ZAP(Stratagene) (T3 primer/EcoRI are at the 5'-end and T7/XhoI
            are at the 3'-end). The library was excised [now in pBluescript
            SK(+)] and normalized (Bonaldo et al 1996 Genome Research
            6:791-806). Library constructed by Thomas Baum (tbaum@iastate.edu),
            Iowa State University, Plant Pathology Department and Jeff
            McDermott (jpmcderm@iastate.edu).
            Seq primer: T3 from Gibco.
FEATURES             Location/Qualifiers
     source          1. .480
                     /organism="Heterodera glycines"
                     /mol_type="mRNA"
                     /db_xref="taxon:51029"
                     /sex="female"
                     /tissue_type="whole organism"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Heterodera glycines virgin female"
                     /note="Vector: pBluescript SK+ (Stratagene); Site_1: XhoI;
                     Site_2: EcoRI; This library was generated by cloning cDNAs
                     directionally into Uni-ZAP(Stratagene) (T3 primer/EcoRI
                     are at the 5'-end and T7/XhoI are at the 3'-end). The
                     library was excised [now in pBluescript SK(+)] and
                     normalized (Bonaldo et al 1996 Genome Research 6:791-806).
                     Library constructed by Thomas Baum (tbaum@iastate.edu),
                     Iowa State University, Plant Pathology Department and Jeff
                     McDermott (jpmcderm@iastate.edu)."
ORIGIN

  Query Match            21.3%;  Score 185.8;  DB 6;  Length 480;
  Best Local Similarity  65.1%;  Pred. No. 6.1e-43;
  Matches  274;  Conservative    0;  Mismatches  147;  Indels    0;  Gaps    0;

Qy          44 CAAAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAG 103
               |||| ||| | |||||| | || || || || ||||| |||||     || ||||| || |
Db          46 CAAAGCCTTTGCTCGAGTTGTATGTGAAGGCCTCAGGGATTGACAGCCGACGCATCGGTG 105

Qy         104 CCGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTG 163
               |      | || || || ||||| ||||||||   | || ||||||| ||  | ||   ||
Db         106 CATGCTTATTTTGCCAAGAATTTTGGATGGAACTCTACGCTCTTTACGAAGTGGGCTGTG 165

Qy         164 CACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCG 223
               || || || || || ||||| || || ||||| |||| ||  || ||| |||||    |
Db         166 TTCGTGTTGAGGTAAAAACTGTGAATGTCAATTCGGAAGGATACAAAAAGCACTTTTCGG 225

Qy         224 GAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAG 283
               ||||||| || || || ||| | ||| | || ||||     | || |  | || ||    |
Db         226 GAGCACAGCCGCCAATAATGGTGGAACAAGACAAAGGCACGGCAAATGCGGACAACACTG 285

Qy         284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
               | || ||  |||| || || ||  |     || ||||||| |||| ||| | || || ||||
Db         286 ACATCGAGCGACGCATTTTCCAACTTTGCAACGAATTCCATGTGCCATTGTTCGAGAAGG 345

Qy         344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
               ||||    |  |  |||| ||  || | ||||||||||||||||||| | || || | ||
Db         346 ATCCGGAAGTGGCAAAGACCATCCAGGATTTGTACAGGAACTTCAAAATATTTCTCCAAG 405

Qy         404 CAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCAC 463
               | |||    | || || || || ||  || | ||       ||||||  || ||  | |
Db         406 CCAAAACGAATTTGGACAAAGGCAAGCAGCAACCCAGCTCCATTGAAGGACTGCCGCCCC 465

Qy         464 A 464
               |
Db         466 A 466


RESULT 11
BI741942
LOCUS       BI741942                396 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  kt83b07.y1 Strongyloides ratti L1 pAMP1 v3 Chiapelli McCarter
            Strongyloides ratti cDNA 5' similar to WP:CE25711 Y105E8E.Y
            ;contains element MER19 repetitive element ;, mRNA sequence.
ACCESSION   BI741942
VERSION     BI741942.1  GI:15742898
KEYWORDS    EST.
SOURCE      Strongyloides ratti
  ORGANISM  Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
            Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
            Putative full length read
            The vector to vector length is 523
            Seq primer: -40RP from Gibco
            High quality sequence stop: 322.
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Strongyloides ratti"
                     /mol_type="mRNA"
                     /db_xref="taxon:34506"
                     /dev_stage="L1"
                     /lab_host="DH10B"
                     /clone_lib="Strongyloides ratti L1 pAMP1 v3 Chiapelli
                     McCarter"
                     /note="Vector: pAMP1 (Gibco); The library was constructed
                     by Brandi Chiapelli and Dr. James  McCarter at Washington
                     University, St. Louis. The cDNA was  made by using
                     Dynabead oligo-dT priming (Dynal). PCR based  library
                     using a modified protocol from the SMART PCR cDNA
                     Synthesis Kit from Clontech. Directionally cloned into the
                     UDG sites of pAMP1. Nematodes were provided by Dr. Mark
                     Viney  of Bristol, UK."
ORIGIN

  Query Match            20.2%;  Score 176.4;  DB 4;  Length 396;
  Best Local Similarity  66.1%;  Pred. No. 3.3e-40;
  Matches  255;  Conservative    0;  Mismatches  131;  Indels    0;  Gaps    0;

Qy         448 GAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAAT 507
               |||| ||| |||  |    ||  || | |     |    | | |  || ||| |  |
Db           7 GAAGGTCTACCACAATCTGTTGTAGCTTGTCGTACAAAATTGTTAGAACAATTGGCATCA 66
```

```
Qy        508 ATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAA 567
              |||||  | ||| || |  | |||   ||  ||||| |  |    | | ||| |   ||||
Db         67 ATTGACAATTTGTTAGCACAAAGAGGTTCAAGATATTTGTTATCACAAAGTTTATGTGAA 126

Qy        568 TATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTT 627
              ||||| |||||  | |||||| |  | ||||||||||||||| ||||      |||  |
Db        127 TATGATTGTGAGTTAATGCCAAGATTACATCATATTCGAATTGTTGGTCAAAGATTATTA 186

Qy        628 GGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGT 687
              |  || ||||||||||||||| ||| || | |||  |||||| || ||||| ||   |
Db        187 GATTTTGAAATTCCACATAATTTTACTTATTTATGGAATTATATTCTTACTGCCTATAGA 246

Qy        688 ACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAA 747
              ||||| || |||||||| ||||| || || ||||| || ||   | || || ||||| |||
Db        247 ACAGCTGCTTTTATTGAAAGTTGCCCAGCTGATCAAGATATCTTACACCATTATAAGGAA 306

Qy        748 CAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAACAAAAACGCACACA 807
              ||| | ||| | || || ||||||||||||||  | ||| | ||||||||||  ||||||
Db        307 CAATTAAATATTTTTACCAATCAACGTGAAACATTACAAGCACCAACAAAAAATCACACA 366

Qy        808 ATTCCGGAAAAAGTGCTATCGGATAT 833
              |||||  |    ||| ||   || ||
Db        367 ATTCCAAAGGCGGTGTTAGATGAAAT 392
```

RESULT 12
CK349229
LOCUS CK349229 450 bp mRNA linear EST 22-DEC-2003
DEFINITION hggfha10D08 Gland Cell LD PCR cDNA Library Heterodera glycines cDNA, mRNA sequence.
ACCESSION CK349229
VERSION CK349229.1 GI:40304842
KEYWORDS EST.
SOURCE Heterodera glycines
ORGANISM Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
REFERENCE 1 (bases 1 to 450)
AUTHORS Gao,B., Allen,R., Maier,T., Davis,E.L., Baum,T.J. and Hussey,R.S.
TITLE Identification of putative parasitism genes expressed in the esophageal gland cells of the soybean cyst nematode, Heterodera glycines
JOURNAL Mol. Plant Microbe Interact. 14 (10), 1247-1254 (2001)
COMMENT Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: trmaier@iastate.edu
Heterodera glycines Gland Cell LD PCR cDNA Library, single pass sequence.
FEATURES Location/Qualifiers
source 1. .450
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell LD PCR cDNA Library"
/note="Organ: gland cell; Vector: pGEM-T Easy"
ORIGIN

Query Match 19.5%; Score 170.2; DB 7; Length 450;
Best Local Similarity 65.9%; Pred. No. 2.2e-38;
Matches 247; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```
Qy         44 CAAAACCTCTTCTCGAGCTCTACGTAAAAGCGTCAGGAATTGATGCTCGCCGCATTGGAG 103
              |||| ||| | |||||| | || || || || ||||| |||||    || ||||| || |
Db         58 CAAAGCCTTTGCTCGAGTTGTATGTGAAGGCCTCAGGGATTGACAGCCGACGCATCGGTG 117

Qy        104 CCGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTCTTTATGAGATTGGAGTTG 163
              |     | || || || ||||| |||||||| | || |||||||| ||  | ||   ||
Db        118 CATGCTTATTTTGCCAAGAATTTTGGATGGAACTCTACGCTCTTTACGAAGTGGGCTGTG 177

Qy        164 CACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACTTTCTCG 223
              || || || || || ||||| || ||  |||| |||| ||  || ||| |||||    |
Db        178 TTCGTGTTGAGGTAAAAACTGTGAATGTCGATTCGGAAGGATACAAAAAGCACTTTTCGG 237

Qy        224 GAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATAATCGAG 283
              ||||||| || || || ||| | ||| | || ||||     | || |  | || ||     |
Db        238 GAGCACAGCCGCCAATAATGGTGGAACAAGACAAAGGCACGGCAAATGCGGACAACACTG 297

Qy        284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
              | || ||  |||| || || ||  |    ||  ||||| |||| ||| | || || ||||
Db        298 ACATCGAGCGACGCATTTTCCAACTTTGCAACCAATTCCATGTGCCATTGTTCGAGAAGG 357

Qy        344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
              ||||    |  |  |||| ||  || | ||||||||||||||||| | || || | ||
Db        358 ATCCGGAAGTGGCAAAGACCATCCAGGATTTGTACAGGAACTTCAAAATATTTCTCCAAG 417

Qy        404 CAAAAGTAGAGTTCG 418
              | |||    | || |
Db        418 CCAAAACGAATTTTG 432
```

RESULT 13
CA868923
LOCUS CA868923 567 bp mRNA linear EST 20-DEC-2002
DEFINITION pw91g06.y1 Haemonchus contortus intestine pAMP1 v1 Haemonchus contortus cDNA 5' similar to WP:CE25711 Y105E8E.Y ;, mRNA sequence.
ACCESSION CA868923
VERSION CA868923.1 GI:27320472
KEYWORDS EST.
SOURCE Haemonchus contortus
ORGANISM Haemonchus contortus
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida; Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
REFERENCE 1 (bases 1 to 567)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1. Intestinal RNA was provided by Dr. Douglas Jasmer of Washington State University (djasmer@vetmed.wsu.edu).
Seq primer: -40RP from Gibco
High quality sequence stop: 410.
FEATURES Location/Qualifiers
source 1. .567
/organism="Haemonchus contortus"
/mol_type="mRNA"
/db_xref="taxon:6289"
/lab_host="DH10B"
/clone_lib="Haemonchus contortus intestine pAMP1 v1"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI; The library was constructed by Claire Murphy and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based

library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1. Intestinal RNA was provided by Dr. Douglas Jasmer of Washington State University (djasmer@vetmed.wsu.edu)."

ORIGIN

Query Match 17.5%; Score 152.6; DB 6; Length 567;
Best Local Similarity 69.8%; Pred. No. 3.4e-33;
Matches 206; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
Qy        559 ATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTG 618
              ||||| |||||||| |||||||| |||||  | || |||||||| || || || ||    |
Db          4 ATGACGGAATATGATTGTGAACTTATGCCCAGACTACATCATATGCGTATCATCGGCCAG 63

Qy        619 TCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACT 678
              | | ||    |||||| | || |  ||||| ||  |  | |||  |||  | || ||
Db         64 AGAATGCTCAATTTCGATTTCCCTCGCAATTTAACCTACTTGTGGAATTACGTACTGACC 123

Qy        679 GCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCAC 738
              || || || || || || || ||||||| ||||| |||||||| || ||  | |||||
Db        124 GCCTATCGAACGGCCGCCTTCATTGAGAGCTGTCCGGCCGATCAAGATATACTGCATCAT 183

Qy        739 TATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAACAAAA 798
              ||||| || ||| | ||| |||||||||||||| ||||| || || || ||||| || ||
Db        184 TATAAGGAGCAACTAAATATGTTCACAAATCAGCGTGAGACGCTACAGTCGCCGACGAAG 243

Qy        799 ACGCACACAATTCCGGAAAAAGTGCTATCGGATATTCGTGTTAAAGGACTTGCTC 853
              || || || |||||||||| | ||| |   ||| ||  |    |  || || | ||
Db        244 ACACATACCATTCCGGAAGAGGTGTTGATGGACATCAGAAGAAGTGGGCTGGATC 298
```

RESULT 14
BG225023
LOCUS BG225023 426 bp mRNA linear EST 09-MAY-2001
DEFINITION kp61c08.y1 TBN95TM-SSFH Strongyloides stercoralis cDNA 5' similar to SW:CLI2_HUMAN O15247 CHLORIDE INTRACELLULAR CHANNEL PROTEIN 2 ;, mRNA sequence.
ACCESSION BG225023
VERSION BG225023.1 GI:12712578
KEYWORDS EST.
SOURCE Strongyloides stercoralis
ORGANISM Strongyloides stercoralis
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE 1 (bases 1 to 426)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Dr. Thomas Nutman and colleagues of NIAID, NIH (tnutman@nih.gov). DNA Sequencing by: Washington University Genome Sequencing Center St. Louis.
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
FEATURES Location/Qualifiers
source 1. .426
/organism="Strongyloides stercoralis"
/mol_type="mRNA"
/strain="Filariform larvae obtained from humans"
/db_xref="taxon:6248"
/lab_host="XL-1 Blue MRF' (Stratagene)"
/clone_lib="TBN95TM-SSFH"
/note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1: EcoRI; Site_2: XhoI; mRNA was purified from 4 x 10E5 filariform larvae which had been isolated from infected humans. cDNA was constructed and, using adaptors, was cloned unidirectionally into the vector from the EcoRI site to the XhoI site. The library has an unamplified titer of 1.5 x 10E6 pfu/ml and an amplified, undiluted titer of 7 x 10E9 pfu/ml. The average insert size of the unamplifed library is 975 bp (range, 500-1500)."

ORIGIN

Query Match 16.8%; Score 146.8; DB 4; Length 426;
Best Local Similarity 70.5%; Pred. No. 1.6e-31;
Matches 196; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```
Qy        575 GTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCG 634
              ||||| |||||||| |  | |||||||| ||||| ||||       | |  | |  || |
Db          1 GTGAATTGATGCCAAGATTACATCATATTAGAATTGTTGGTCAAAGATTATTAGATTTTG 60

Qy        635 ATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAG 694
              | |||||||||||||| ||  || | |||  |||||| || ||||| ||  | || || |
Db         61 AAATTCCACATAATTTTACCTATTTATGGAATTATATTCTTACTGCCTATAGAACTGCTG 120

Qy        695 CATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGA 754
              | || ||||| ||||| || || ||||| || ||| | || || || || |||||| | |
Db        121 CTTTCATTGAAAGTTGCCCAGCTGATCAAGATATTTTACACCATTACAAGGAACAATTAA 180

Qy        755 ATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAACAAAAACGCACACAATTCCGG 814
              || | || ||||||||||||||||||  | ||| | ||||||||||  |||||||||| |
Db        181 ATATTTTTACAAATCAACGTGAAACATTACAAGCACCAACAAAAAATCACACAATTCCTG 240

Qy        815 AAAAAGTGCTATCGGATATTCGTGTTAAAGGACTTGCT 852
              |   |||| ||   || ||  |       |  ||||| |
Db        241 AGGCAGTGTTAGATGAAATAAGGCGATTAAAACTTGAT 278
```

RESULT 15
CV508408
LOCUS CV508408 324 bp mRNA linear EST 05-OCT-2004
DEFINITION kc71g05.y1 Xiphinema index CSEQDL01 Xiphinema index cDNA 5' similar to WP:CE25711 Y105E8E.Y ;, mRNA sequence.
ACCESSION CV508408
VERSION CV508408.1 GI:53806521
KEYWORDS EST.
SOURCE Xiphinema index
ORGANISM Xiphinema index
Eukaryota; Metazoa; Nematoda; Enoplea; Dorylaimida; Dorylaimina; Longidoroidea; Longidoridae; Xiphinema.
REFERENCE 1 (bases 1 to 324)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Mitreva M
The Washington Univ. Nematode EST Project, 2004
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library materials provided by: J.T. Jones Library constructed by: Incyte Genomics (for J.T. Jones)
Putative full length read
The vector to vector length is 325
Seq primer: -21UPpot primer.
FEATURES Location/Qualifiers

```
     source          1. .324
                     /organism="Xiphinema index"
                     /mol_type="mRNA"
                     /db_xref="taxon:46003"
                     /sex="mixed"
                     /dev_stage="mixed"
                     /lab_host="GC10 (only ligation mix supplied)"
                     /clone_lib="Xiphinema index CSEQDL01"
                     /note="Vector: pBluescript SK+; Site_1: 5': NotI; Site_2:
                     3': EcoRI; The library was made from mRNA isolated from
                     total RNA using oligo dT cellulose. Total RNA was
                     generated from mixed stage X. index extracted from soil
                     around fig (Ficus carica) roots. Nematodes were cleaned by
                     passing through a series of sieves.  Pure X. index were
                     obtained by hand sorting of nematodes. The library was
                     constructed by Incyte Genomics by cloning cDNA
                     directionally into in the pBluescript (SK+) vector using
                     NotI and EcoRI sites. The library was provided by John T.
                     Jones (jjones@scri.sari.ac.uk) at SCRI, Nematology
                     Department. Sequencing by: Washington University Genome
                     Sequencing Center, St. Louis, MO."
ORIGIN

  Query Match            15.9%;  Score 138.6;  DB 7;  Length 324;
  Best Local Similarity  66.7%;  Pred. No. 3.6e-29;
  Matches  198;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy        536 CTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTC 595
              | || |||||  | ||       ||||| ||||| || || ||| ||||||||||| || |
Db         26 CACGTTATCTTTTGGGTTCGTCCATGACGGAATACGATTGCGAATTGATGCCACGCCTGC 85

Qy        596 ATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTC 655
              ||||||||||||||    || |    | |    ||  | || |||||||| ||| |   |
Db         86 ATCATATTCGAATCGCCGGTTCAAAATTATGCGGTGTGGAAATTCCACACAATCTGGTTT 145

Qy        656 ATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCG 715
              || | |||  |||||| || ||  | || || || || |||||||| |||   || || |
Db        146 ATTTGTGGAATTATATGCTGACGACGTATCGAACTGCCGCATTTATCGAGTCGTGCCCGG 205

Qy        716 CCGATCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTG 775
              | ||||| |||||  |||   || || |||||||| ||||  |   |      ||||| |
Db        206 CTGATCAAGACATCATTTTCCATTACAAAGAACAATTGAACATTCCCCTGCGACAACGCG 265

Qy        776 AAACCCTCCAATCGCCAACAAAAACGCACACAATTCCGGAAAAAGTGCTATCGGATA 832
              |||  || |||| ||| |||||||||||| || ||||| ||| | ||  |  | || |
Db        266 AAAGTCTTCAATTGCCTACAAAAACGCATACGATTCCCGAAGATGTCTTGGCCGAAA 322


Search completed: March 10, 2005, 02:20:41
Job time : 3566 secs
```

This page blank (uspto)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 23:54:49 ; Search time 194 Seconds
(without alignments)
7363.245 Million cell updates/sec

Title: US-10-612-379-1
Perfect score: 873
Sequence: 1 atggcagaagcttaccagat.....ccgatgttaatgttcattaa 873

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	76.4	8.8	1090	4	US-09-270-767-26167	Sequence 26167, A
	2	76.4	8.8	1880	4	US-09-270-767-10712	Sequence 10712, A
c	3	47.4	5.4	7218	1	US-08-232-463-14	Sequence 14, Appl
	4	46.4	5.3	819	3	US-08-792-014-2	Sequence 2, Appli
	5	46.4	5.3	819	3	US-09-443-948-2	Sequence 2, Appli
	6	46.4	5.3	819	4	US-09-690-196-2	Sequence 2, Appli
	7	43.4	5.0	1652	4	US-09-620-312D-815	Sequence 815, App
	8	43.4	5.0	1689	4	US-09-949-016-5780	Sequence 5780, Ap
	9	41	4.7	1068	4	US-09-710-794-3	Sequence 3, Appli
	10	39.8	4.6	601	4	US-09-949-016-107953	Sequence 107953,
	11	39.8	4.6	601	4	US-09-949-016-107954	Sequence 107954,
c	12	39.8	4.6	390890	4	US-09-949-016-14720	Sequence 14720, A
	13	39.2	4.5	1529	4	US-09-533-029-49	Sequence 49, Appl
	14	38.4	4.4	1141	4	US-09-806-708B-22	Sequence 22, Appl
	15	38.4	4.4	253345	4	US-09-949-016-12656	Sequence 12656, A
	16	38.4	4.4	253364	4	US-09-949-016-13639	Sequence 13639, A
	17	37.8	4.3	832	4	US-09-621-976-2813	Sequence 2813, Ap
c	18	37	4.2	14066	4	US-09-601-198-56	Sequence 56, Appl
c	19	36.8	4.2	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c	20	36.8	4.2	1664976	4	US-09-692-570-1	Sequence 1, Appli
	21	36	4.1	42381	4	US-09-949-016-12012	Sequence 12012, A
	22	36	4.1	168394	4	US-09-949-016-13002	Sequence 13002, A
	23	36	4.1	183770	4	US-09-949-016-15494	Sequence 15494, A
	24	35.8	4.1	95561	4	US-09-949-016-12768	Sequence 12768, A
	25	35.8	4.1	95561	4	US-09-949-016-13306	Sequence 13306, A
	26	35.8	4.1	95561	4	US-09-949-016-13307	Sequence 13307, A
c	27	35.6	4.1	601	4	US-09-949-016-86515	Sequence 86515, A
	28	35.6	4.1	61462	4	US-09-949-016-17522	Sequence 17522, A
c	29	35.6	4.1	183112	4	US-09-949-016-14184	Sequence 14184, A
c	30	35.4	4.1	601	4	US-09-949-016-141596	Sequence 141596,
c	31	35.4	4.1	832	4	US-09-621-976-2813	Sequence 2813, Ap
	32	35.4	4.1	236964	4	US-09-949-016-15753	Sequence 15753, A
c	33	35.4	4.1	250715	4	US-09-949-016-13294	Sequence 13294, A
c	34	35	4.0	601	4	US-09-949-016-141597	Sequence 141597,
	35	35	4.0	1497	3	US-09-232-468A-17	Sequence 17, Appl
	36	35	4.0	1497	4	US-09-784-984B-14	Sequence 14, Appl
	37	34.6	4.0	696	4	US-09-583-110-1334	Sequence 1334, Ap
	38	34.6	4.0	789	4	US-09-107-433-176	Sequence 176, App
	39	34.6	4.0	921	4	US-09-270-767-8688	Sequence 8688, Ap
	40	34.6	4.0	921	4	US-09-270-767-23970	Sequence 23970, A
	41	34.2	3.9	1086	4	US-09-248-796A-59	Sequence 59, Appl
	42	34.2	3.9	2082	4	US-09-248-796A-2564	Sequence 2564, Ap
c	43	34.2	3.9	343352	4	US-09-949-016-13498	Sequence 13498, A
	44	34	3.9	55806	4	US-09-949-016-15605	Sequence 15605, A
	45	33.8	3.9	601	4	US-09-949-016-37936	Sequence 37936, A

ALIGNMENTS

RESULT 1
US-09-270-767-26167
; Sequence 26167, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26167
; LENGTH: 1090
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-26167

Query Match 8.8%; Score 76.4; DB 4; Length 1090;
Best Local Similarity 52.5%; Pred. No. 9.2e-14;
Matches 167; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

```
Qy      495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
            ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||   ||  |
Db       28 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 87

Qy      555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
            ||  |||      | |||||||| |||||||| || || || || || ||  |    ||
Db       88 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 147

Qy      615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
              ||  || | |  || || || ||     | || ||    || |||   || ||
Db      148 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 207

Qy      675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                ||||      |  || || |   |     || || ||||| ||||||||||  |
Db      208 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 267

Qy      735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
            |||||| ||    |||  || |||    |  ||  | |  ||    ||  |  |||| ||
Db      268 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 327

Qy      795 AAAAACGCACACAATTCC 812
              ||      |||||
Db      328 GTTTACCACATACATTCC 345
```

RESULT 2
US-09-270-767-10712
; Sequence 10712, Application US/09270767

```
; Patent No. 6703491
; GENERAL INFORMATION:
;  APPLICANT: Homburger et al.
;  TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;  FILE REFERENCE: File Reference: 7326-094
;  CURRENT APPLICATION NUMBER: US/09/270,767
;  CURRENT FILING DATE: 1999-03-17
;  NUMBER OF SEQ ID NOS: 62517
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10712
;   LENGTH: 1880
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-10712

  Query Match             8.8%;  Score 76.4;  DB 4;  Length 1880;
  Best Local Similarity   52.5%;  Pred. No. 1.3e-13;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    ||  |
Db        818 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 877

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              ||  |||       |  |||||||| |||||||| || || || || || || |    ||
Db        878 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 937

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||   || | |  || || || ||      | || ||    || |||   || ||
Db        938 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 997

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                ||||      |  || || |   |    || || ||||| ||||||||||  |
Db        998 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 1057

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||    |||  || |||    |  ||  | |  ||    ||  |  |||| ||
Db       1058 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 1117

Qy        795 AAAAACGCACACAATTCC 812
                ||       |||||
Db       1118 GTTTACCACATACATTCC 1135


RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             5.4%;  Score 47.4;  DB 1;  Length 7218;
  Best Local Similarity   7.4%;  Pred. No. 0.00047;
  Matches   27;  Conservative  187;  Mismatches  153;  Indels    0;  Gaps    0;

Qy         86 ATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAGGAATTCTGGATGGAGTTGTATGCTC 145
              |||   ||  || || |         | | ||  || | ||   ||   ||  | ||
Db       1497 ATGTAGGCATCACTGTAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAAGAATTTGG 1438

Qy        146 TTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCAT 205
              |  : ::::  ::::  : : :::  :::: :::::  :  :: : :::    :::: :
Db       1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378

Qy        206 TTAAGAAGAACTTTCTCGGAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGA 265
               ::::::::       :::: : ::  :  ::  : ::  ::::::::::::::  ::
Db       1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318

Qy        266 CATACACTGATAATCGAGAGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATG 325
               : : :  :: ::  ::::::  :::::: :::       :   :: :::::::   :: :
Db       1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258

Qy        326 TTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACT 385
              :      ::::::::   :   :  :::::::::: ::::::   : : :::::
Db       1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198

Qy        386 TCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAG 445
               :::  :     : ::: ::::: ::::   :: ::::::::::::::::  :  :::::
Db       1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138

Qy        446 TTGAAGA 452
               :::::
Db       1137 RRRRRRR 1131


RESULT 4
US-08-792-014-2
; Sequence 2, Application US/08792014
; Patent No. 6063594
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN ANION CHANNEL
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
```

```
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/792,014
;       FILING DATE:  Herewith
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Billings, Lucy J.
;       REGISTRATION NUMBER:  36,749
;       REFERENCE/DOCKET NUMBER:  PF-0206 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-855-0555
;       TELEFAX:  415-845-4166
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  819 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  Consensus
;       CLONE:  Consensus
US-08-792-014-2

  Query Match             5.3%;  Score 46.4;  DB 3;  Length 819;
  Best Local Similarity   50.0%;  Pred. No. 0.00029;
  Matches  116;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   | ||| ||  | ||| ||||    || |||
Db        546 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 605

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     | | |       |  ||     || |||||| || |  |  | |||    |
Db        606 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTTCAAAAGAAATGACTGGCAT 665

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||||| |||  |  |  ||  |   | |  ||||||   ||
Db        666 CTGGAGATACCTAACTAATGCATACAGTAGGGACGGGTTCACCAATACCTGTCCCAGTGA 725

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||         ||||  ||   |   ||     |||| ||  ||
Db        726 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 777


RESULT 5
US-09-443-948-2
; Sequence 2, Application US/09443948
; Patent No. 6228616
;   GENERAL INFORMATION:
;       APPLICANT: Bandman, Olga
;                  Goli, Surya K.
;       TITLE OF INVENTION: NOVEL HUMAN ANION CHANNEL
;       NUMBER OF SEQUENCES: 4
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Drive
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/443,948
;            FILING DATE: 19-No. 6228616-1999
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/792,014
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0206 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 415-855-0555
;            TELEFAX: 415-845-4166
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 819 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: Consensus
;            CLONE: Consensus
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-443-948-2

  Query Match             5.3%;  Score 46.4;  DB 3;  Length 819;
  Best Local Similarity   50.0%;  Pred. No. 0.00029;
  Matches  116;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   | ||| ||  | ||| ||||    || |||
Db        546 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 605

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     | | |       |  ||     || |||||| || |  |  | |||    |
Db        606 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTTCAAAAGAAATGACTGGCAT 665

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||||| |||  |  |  ||  |   | |  ||||||   ||
Db        666 CTGGAGATACCTAACTAATGCATACAGTAGGGACGGGTTCACCAATACCTGTCCCAGTGA 725

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||         ||||  ||   |   ||     |||| ||  ||
Db        726 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 777


RESULT 6
US-09-690-196-2
; Sequence 2, Application US/09690196
; Patent No. 6503733
;   GENERAL INFORMATION:
;       APPLICANT: Bandman, Olga
;                  Goli, Surya K.
;       TITLE OF INVENTION: NOVEL HUMAN ANION CHANNEL
;       NUMBER OF SEQUENCES: 4
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Drive
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/690,196
;            FILING DATE: 16-Oct-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/792,014
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
```

```
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0206 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415-855-0555
;             TELEFAX: 415-845-4166
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 2:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 819 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: Consensus
;             CLONE: Consensus
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-690-196-2

  Query Match             5.3%;  Score 46.4;  DB 4;  Length 819;
  Best Local Similarity   50.0%;  Pred. No. 0.00029;
  Matches  116;  Conservative    0;  Mismatches  116;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||    | ||| ||  | ||| ||||    || |||
Db        546 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 605

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     | | |        |  ||     || |||||| || |  |  | |||    |
Db        606 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTTCAAAAGAAATGACTGGCAT 665

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||||| |||  |  |  || |   | |  ||||||   ||
Db        666 CTGGAGATACCTAACTAATGCATACAGTAGGGACGGGTTCACCAATACCTGTCCCAGTGA 725

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||         ||||  ||   |   ||     |||| ||  ||
Db        726 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 777


RESULT 7
US-09-620-312D-815
; Sequence 815, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 815
;   LENGTH: 1652
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (232)..(975)
US-09-620-312D-815

  Query Match             5.0%;  Score 43.4;  DB 4;  Length 1652;
  Best Local Similarity   52.5%;  Pred. No. 0.0038;
  Matches   95;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy        549 TGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAAT 608
              |||  ||    | ||      ||| |||    || | ||    ||  |   ||||  | |
Db        756 TGGGGACCAGCTAACACTGGCTGATTGTAGCTTGTTACCCAAGCTGAACATTATTAAAGT 815

Qy        609 TATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTA 668
              |  ||    |  | || ||  || || ||||||    | ||| |      |||||  |||
Db        816 TGCTGCCAAGAAATATCGTGACTTTGACATTCCAGCAGAATTCTCAGGAGTCTGGCGTTA 875

Qy        669 TATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACAT 728
              | ||| || ||| ||       | || |||||   | |  |||||  |  ||  | || ||
Db        876 TCTCCACAATGCCTATGCCCGTGAAGAATTTACCCACACGTGTCCTGAAGACAAAGAAAT 935

Qy        729 T 729
              |
Db        936 T 936


RESULT 8
US-09-949-016-5780
; Sequence 5780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5780
;   LENGTH: 1689
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-5780

  Query Match             5.0%;  Score 43.4;  DB 4;  Length 1689;
  Best Local Similarity   52.5%;  Pred. No. 0.0039;
  Matches   95;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy        549 TGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAAT 608
              |||  ||    | ||      ||| |||    || | ||    ||  |   ||||  | |
Db        746 TGGGGACCAGCTAACACTGGCTGATTGTAGCTTGTTACCCAAGCTGAACATTATTAAAGT 805

Qy        609 TATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTA 668
              |  ||    |  | || ||  || || ||||||    | ||| |      |||||  |||
Db        806 TGCTGCCAAGAAATATCGTGACTTTGACATTCCAGCAGAATTCTCAGGAGTCTGGCGTTA 865

Qy        669 TATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACAT 728
              | ||| || ||| ||       | || |||||   | |  |||||  |  ||  | || ||
Db        866 TCTCCACAATGCCTATGCCCGTGAAGAATTTACCCACACGTGTCCTGAAGACAAAGAAAT 925

Qy        729 T 729
              |
```

```
Db        926 T 926


RESULT 9
US-09-710-794-3
; Sequence 3, Application US/09710794
; Patent No. 6573069
; GENERAL INFORMATION:
;  APPLICANT:  Holloway, James L.
;  APPLICANT:  Gao, Zeren
;  APPLICANT:  Whitmore, Theodore E.
;  TITLE OF INVENTION:  NOVEL CRIB PROTEIN ZMSE1
;  FILE REFERENCE:  99-76
;  CURRENT APPLICATION NUMBER: US/09/710,794
;  CURRENT FILING DATE:  2000-11-09
;  PRIOR APPLICATION NUMBER:  US 60/164,685
;  PRIOR FILING DATE:  1999-11-10
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;    LENGTH: 1068
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Degenerate polynucleotide seuence for human zmse1
;    NAME/KEY: misc_feature
;    LOCATION: (1)...(1068)
;    OTHER INFORMATION: n = A,T,C or G
US-09-710-794-3

  Query Match             4.7%;  Score 41;  DB 4;  Length 1068;
  Best Local Similarity   25.7%;  Pred. No. 0.017;
  Matches   76;  Conservative   59;  Mismatches  160;  Indels    1;  Gaps    1;

Qy        158 GAGTTGCACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAAGAAGAACT 217
              ||::  :  ||   : |:   ::    :: :: :: |  :  :  :  :| :: :| |
Db        187 GARWSNYTNGAYGARCARCCNWSNWSNWSNWSNWSNAARMGNWSNYTNYTNWSNMGNAAR 246

Qy        218 TTCTCGGAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGACATACACTGATA 277
              ||:   |     || :   :   ::: | |  :|     |::|   :|| :     ||
Db        247 TTYMGNGGNWSNAARMGNWSNCARWSNGTNACNMGNGGNGARMGNGARCARMGNGAYATG 306

Qy        278 ATCGAGAGATTGAAGGACGGATCTTTCATT-TGGCAAAGGAATTCAATGTTCCACTCTTT 336
              | |  :: :| :  ||: :  | :|  |: |   :|| |   | :   | || |  :|
Db        307 YTNGGNWSNYTNMGNGAYWSNGCNYTNTTYGTNAARAAYGCNATGWSNYTNCCNCARYTN 366

Qy        337 GAAAAGGATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTC 396
              |  |: |   : | || ||:||: | |     || :|    | ::  :| :     :
Db        367 AAYGARAARGARGCNGCNGARAARGGNACNWSNAARYTNCCNAARWSNYTNWSNWSNWSN 426

Qy        397 CTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGTTGAAGA 452
              |       :||:|   |   :|   |:||     | ||:  : | |    ||:||
Db        427 CCNGTNAARAARGCNAAYGAYGGNGARGGNGGNGAYGARGARGCNGGNACNGARGA 482


RESULT 10
US-09-949-016-107953
; Sequence 107953, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107953
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-107953

  Query Match             4.6%;  Score 39.8;  DB 4;  Length 601;
  Best Local Similarity   52.0%;  Pred. No. 0.03;
  Matches   89;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy        324 TGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAA 383
              |||   |||||  ||||||  | |  |       ||||||  |   |  ||    |||| ||
Db        367 TGTAAAACTCTGGGAAAAGTAACATTTAAGGAAGAAGAACAAGAAAGACAGTAACAGAAA 426

Qy        384 CTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAG 443
              | |    ||||||||  ||  |||     |   | |      |||    |  |  ||
Db        427 ATACTCTATGTTCCTGTGAATAAAGACTAATACCTAGGTATAAAATTCAGGACAGTCTCA 486

Qy        444 AGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGA 494
              |||| ||  | |||||||  || ||  | | |   | || ||    || |||
Db        487 AGTTTAATAACTTCCAGTCCACATATATGATGTTTCCATTCTGTGCTTTGA 537


RESULT 11
US-09-949-016-107954
; Sequence 107954, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 107954
;    LENGTH: 601
;    TYPE: DNA
;    ORGANISM: Human
US-09-949-016-107954

  Query Match             4.6%;  Score 39.8;  DB 4;  Length 601;
  Best Local Similarity   52.0%;  Pred. No. 0.03;
  Matches   89;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy        324 TGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAA 383
              |||   |||||  ||||||  | |  |       ||||||  |   |  ||    |||| ||
Db        368 TGTAAAACTCTGGGAAAAGTAACATTTAAGGAAGAAGAACAAGAAAGACAGTAACAGAAA 427

Qy        384 CTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAG 443
              | |    ||||||||  ||  |||     |   | |      |||    |  |  ||
Db        428 ATACTCTATGTTCCTGTGAATAAAGACTAATACCTAGGTATAAAATTCAGGACAGTCTCA 487

Qy        444 AGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGA 494
              |||| ||  | |||||||  || ||  | | |   | || ||    || |||
Db        488 AGTTTAATAACTTCCAGTCCACATATATGATGTTTCCATTCTGTGCTTTGA 538


RESULT 12
US-09-949-016-14720/c
; Sequence 14720, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14720
;   LENGTH: 390890
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(390890)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14720

  Query Match             4.6%;  Score 39.8;  DB 4;  Length 390890;
  Best Local Similarity   52.0%;  Pred. No. 1.2;
  Matches   89;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy        324 TGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAA 383
              |||   |||||  |||||| | |  |       ||||||  |   |  ||    |||| ||
Db     314235 TGTAAAACTCTGGGAAAAGTAACATTTAAGGAAGAAGAACAAGAAAGACAGTAACAGAAA 314176

Qy        384 CTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAG 443
              | |    |||||||| ||  |||      |   | |       |||    |  |  ||
Db     314175 ATACTCTATGTTCCTGTGAATAAAGACTAATACCTAGGTATAAAATTCAGGACAGTCTCA 314116

Qy        444 AGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGA 494
              |||| || | |||||||  || ||  | | |   | || ||    || |||
Db     314115 AGTTTAATAACTTCCAGTCCACATATATGATGTTTCCATTCTGTGCTTTGA 314065


RESULT 13
US-09-533-029-49
; Sequence 49, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
;  APPLICANT: Heard, Jacqueline
;  APPLICANT: Broun, Pierre
;  APPLICANT: Riechmann, Jose-Luis
;  APPLICANT: Keddie, James
;  APPLICANT: Pineda, Omaira
;  APPLICANT: Adam, Luc
;  APPLICANT: Samaha, Raymond
;  APPLICANT: Zhang, James
;  APPLICANT: Yu, Guo-Liang
;  APPLICANT: Ratcliffe, Oliver
;  APPLICANT: Pilgrim, Marsha
;  APPLICANT: Jiang, Cai-Zhong
;  APPLICANT: Reuber, Lynne
;  TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
;  FILE REFERENCE: MBI-010
;  CURRENT APPLICATION NUMBER: US/09/533,029
;  CURRENT FILING DATE: 2000-03-22
;  EARLIER APPLICATION NUMBER: 60/125,814
;  EARLIER FILING DATE: 1999-03-23
;  NUMBER OF SEQ ID NOS: 121
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
;   LENGTH: 1529
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: G553
US-09-533-029-49

  Query Match             4.5%;  Score 39.2;  DB 4;  Length 1529;
  Best Local Similarity   58.6%;  Pred. No. 0.08;
  Matches   68;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy        150 TGAGATTGGAGTTGCACGAGTCGAAGTGAAGACTGTCAACGTGAATTCTGAAGCATTTAA 209
              |||| |||  | |  || ||  | ||   | |      | ||||||  |    |  |  |
Db        249 TGAGGTTGATGCTAGACCAGAAGCAGATGATAACAATAGAGTGAATTATACTTCTGTGTA 308

Qy        210 GAAGAACTTTCTCGGAGCACAACCACCGATTATGATTGAAGAGGAAAAAGAGCTGA 265
              || |||| ||| | |||| ||||  ||| ||  | |||  ||||  |||| | ||
Db        309 TAATAACTCTCTTGAAGCAGAACCGTCGAGTAATAATGATCAGGACGAAGACCGGA 364


RESULT 14
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             4.4%;  Score 38.4;  DB 4;  Length 1141;
  Best Local Similarity   8.5%;  Pred. No. 0.12;
  Matches   36;  Conservative  183;  Mismatches  207;  Indels    0;  Gaps    0;

Qy        284 AGATTGAAGGACGGATCTTTCATTTGGCAAAGGAATTCAATGTTCCACTCTTTGAAAAGG 343
              :::|: ::    |  |:|: :::::|:: :  ::: : : ::: ::::::|:::::::::: :
Db        176 DDDTKYHMWNNNGCBTVTWMVRYKTDRDWSBKRMNYGMBWWKNWSYDVTYYWWVWDDMCK 235

Qy        344 ATCCATCCGCTGAGAAGAGAATAGAGAACTTGTACAGGAACTTCAAACTGTTCCTGCGAG 403
              ::: :::    :::   : : || :::      |  :  :: |         : :|:: :
Db        236 RKVRRWVRTRGRMRNYMVAWBTAHRRRYNNGWTBAMAYRRWTMNNNNNNAKAMCKRAKYW 295

Qy        404 CAAAAGTAGAGTTCGATAAGGGAAAAAAGGAGCCATCGAGAGTTGAAGATCTTCCAGCAC 463
              : :|:  :  || : :  :::   :|   :   |  :::: :::||   ::
Db        296 GWNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDHKWWKWSAAMGVYWNNNNNNNW 355

Qy        464 AGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCTAT 523
                : ::   : :::::|::  :     :::: : ::    ::    | |:::    : :
Db        356 TYKKARHBARWDWVWHSAWKKWHANAAHYSRKKWTBYKRKTMVNNNNGTTMWKRMWAWYW 415

Qy        524 CCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGA 583
              ::::::         |   :: : : :::: : ::    |: :|:::: ||       :::
Db        416 KMDMDWBGTYNNNNNGGRTYYGWTKNKKMWTYYKWKANNCKWRAWDHKTCTHNNTTWWKM 475

Qy        584 TGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCAC 643
              : :    : : :     : :  |::|::   :    :|    ::  :: | :: :::|
Db        476 KTYWNNCYWKSMTNGKSHRBAAAVYTWYMWWWRRYAHANNNNWDYWWKACTWYKYBVCSK 535

Qy        644 ATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTG 703
              ::  :    :: : :  :  ::      :   |::: :   :| |:: |: : |::::
Db        536 WWNNYAAWYTKSSWNYTSRYYRWKTNNSWRWRSDTRSMGRANNYARABHYGYKWNTRWWB 595

Qy        704 AGAGTT 709
              ::: ::
Db        596 WSHTWB 601
```

RESULT 15
US-09-949-016-12656
; Sequence 12656, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12656
; LENGTH: 253345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12656

Query Match 4.4%; Score 38.4; DB 4; Length 253345;
Best Local Similarity 57.5%; Pred. No. 2.6;
Matches 69; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

```
Qy        438 ATCGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCA 497
              |||   |||  || || ||  |||| ||||| |||||||  |  || |    || ||   |
Db      17149 ATCATCAGTCTAATATTTTATAGCAAAGATTTAAGTTCATTGAAAACTTCCCTATGGAAA 17208

Qy        498 ACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAACAG 557
              ||||||   || |    || | ||  ||  ||||    || |  || || |||| |
Db      17209 CACTTCCAATCAAGACCTACAGCAAGCCATGAATAAATTAACATTTAAACATGAAAACTG 17268
```

Search completed: March 10, 2005, 02:24:01
Job time : 202 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 22:58:24 ; Search.time 588 Seconds
(without alignments)
8789.004 Million cell updates/sec

Title: US-10-612-379-1
Perfect score: 873
Sequence: 1 atggcagaagcttaccagat.....ccgatgttaatgttcattaa 873

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		76.4	8.8	783	13	ADS96405	Ads96405 Drosophil
2		76.4	8.8	1063	4	ABL17133	Abl17133 Drosophil
3	c	63	7.2	786	3	AAC95377	Aac95377 Cat flea
4		63	7.2	786	3	AAC95376	Aac95376 Cat flea
5	c	63	7.2	2383	3	AAC95375	Aac95375 Cat flea
6		63	7.2	2383	3	AAC95374	Aac95374 Cat flea
7	c	51.2	5.9	2000	8	ADA71938	Ada71938 Rice gene
8		48	5.5	2238	11	ADM02307	Adm02307 Human cDN
9		48	5.5	2887	9	ADA24507	Ada24507 Human cDN
10		48	5.5	4290	8	ACC83320	Acc83320 Chloride
11		48	5.5	4318	4	AAK51503	Aak51503 Human pol
12		48	5.5	4318	4	AAI57882	Aai57882 Human pol
13		48	5.5	4318	8	ACF12844	Acf12844 Human cer
14		48	5.5	4318	13	ADR25580	Adr25580 Breast ca
15		48	5.5	4318	13	ADP23232	Adp23232 PRO polyp
16		48	5.5	4707	11	ADM86762	Adm86762 Human cDN
17		46.4	5.3	819	2	AAV43260	Aav43260 DNA encod
18		46.4	5.3	819	10	ADJ56392	Adj56392 Human cDN
19		46	5.3	4357	4	AAK52487	Aak52487 Human pol
20		46	5.3	4357	4	AAI59668	Aai59668 Human pol
21		46	5.3	4502	5	AAS87686	Aas87686 DNA encod
22		43.4	5.0	371	8	ABX34992	Abx34992 Bovine ES
23		43.4	5.0	1229	4	AAH02899	Aah02899 Human she
24		43.4	5.0	1652	4	AAI58923	Aai58923 Human pol
25		43.4	5.0	1652	5	ADQ99145	Adq99145 DNA encod
26		43.4	5.0	1652	9	ADB48905	Adb48905 Novel hum
27		43.4	5.0	1701	13	ADQ87447	Adq87447 Human tum
28		43.4	5.0	1713	4	AAI60709	Aai60709 Human pol
29		43.4	5.0	1722	5	ABV29180	Abv29180 Human pro
30		43.4	5.0	1722	5	ABV23328	Abv23328 Human pro
31		43.4	5.0	2260	10	ADB47440	Adb47440 Human cDN
32		43.4	5.0	2662	10	AAD47905	Aad47905 Human tra
33		43	4.9	10560	4	ABL17132	Abl17132 Drosophil
34		42.8	4.9	624	12	ADP28695	Adp28695 Human sec
35		42.8	4.9	859	6	ADI16569	Adi16569 Human NOV
36		42.8	4.9	859	12	ADN42223	Adn42223 Human cDN
37		42.8	4.9	2061	12	ADJ92742	Adj92742 Human HCL
38		42.8	4.9	2257	12	ADH22598	Adh22598 cDNA enco
39	c	42.8	4.9	2320	6	ABZ25018	Abz25018 Laminin B
40		42.8	4.9	2335	10	ADB63066	Adb63066 Human cDN
41		42.8	4.9	2430	6	AAD27280	Aad27280 Human tra
42		42.8	4.9	3389	12	ADH22570	Adh22570 cDNA enco
43		42.8	4.9	3524	12	ADP99220	Adp99220 Human tra
44		42.8	4.9	3641	12	ADJ92727	Adj92727 Human int
45		41.4	4.7	2000	8	ADA71938	Ada71938 Rice gene

ALIGNMENTS

RESULT 1
ADS96405
ID ADS96405 standard; cDNA; 783 BP.
XX
AC ADS96405;
XX
DT 02-DEC-2004 (first entry)
XX
DE Drosophila melanogaster protein coding sequence, SEQ ID 26.
XX
KW Insecticide; Antiparasitic; Antihelminthic; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO2004039999-A2.
XX
PD 13-MAY-2004.
XX
PF 08-AUG-2003; 2003WO-US024982.
XX
PR 30-OCT-2002; 2002US-0422377P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Stam L, Kamdar KP, Spana E, Bachmann J;
XX
DR WPI; 2004-376203/35.
DR P-PSDB; ADS96406.
XX
PT Identifying a compound that inhibits the activity of a protein for
PT Drosophila viability for use e.g., as insecticidal agent by expressing in
PT a recombinant host a DNA molecule to produce a protein essential for
PT Drosophila viability.
XX
PS Claim 1; SEQ ID NO 26; 57pp; English.
XX
CC The present invention relates to a method for identifying a compound that
CC inhibits the activity of a protein essential for Drosophila viability.
CC The method comprises: (a) expressing in a recombinant host a DNA sequence
CC encoding a protein essential for Drosophila viability; (b) testing
CC compounds suspected of having the ability to inhibit the activity of the
CC protein expressed in (a); and identifying a compound tested in (b) that
CC inhibits the activity of the protein. The method is useful in identifying

```
CC   a compound that inhibits the activity of a protein essential for
CC   Drosophila viability for use as insecticidal, ectoparasiticidal,
CC   antiparasitic, antihelminthic or acaracidal agent. The present sequence
CC   is the DNA sequence for one such protein essential for Drosophila
CC   viability. Note: The sequence data for this patent did not form part of
CC   the printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 783 BP; 207 A; 215 C; 206 G; 155 T; 0 U; 0 Other;

  Query Match             8.8%;  Score 76.4;  DB 13;  Length 783;
  Best Local Similarity   52.5%;  Pred. No. 3e-12;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    ||  |
Db        447 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 506

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              ||  |||      |  |||||||| |||||||| ||.|| || || || ||  |    ||
Db        507 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 566

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||   || | |  || || || ||     | || ||    || |||   || ||
Db        567 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 626

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                 ||||      |  || || |   |    || || ||||| ||||||||||| |
Db        627 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 686

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||    |||  || |||    |  ||  | |  ||    ||  |  |||| ||
Db        687 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 746

Qy        795 AAAAACGCACACAATTCC 812
                ||       |||||
Db        747 GTTTACCACATACATTCC 764


RESULT 2
ABL17133
ID   ABL17133 standard; DNA; 1063 BP.
XX
AC   ABL17133;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 2872.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 2872; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1063 BP; 295 A; 272 C; 271 G; 225 T; 0 U; 0 Other;

  Query Match             8.8%;  Score 76.4;  DB 4;  Length 1063;
  Best Local Similarity   52.5%;  Pred. No. 3.4e-12;
  Matches  167;  Conservative    0;  Mismatches  151;  Indels    0;  Gaps    0;

Qy        495 GCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTGGAAA 554
              ||| ||   ||| ||  |  |    || |||| | | ||  | || |  ||    ||  |
Db        727 GCATCTGCGCAAGATCAACGATCATCTGTCCGCGCGCAACACGCGCTTCCTCACCGGCGA 786

Qy        555 CAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTATTGG 614
              ||  |||      |  |||||||| |||||||| || || || || || ||  |    ||
Db        787 CACCATGTGCTGCTTCGACTGTGAGCTGATGCCGCGCCTGCAGCACATCCGTGTGGCGGG 846

Qy        615 ATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATATCCT 674
                ||   || | |  || || || ||     | || ||    || |||   || ||
Db        847 CAAGTACTTTGTCGACTTTGAAATCCCGACGCACTTGACGGCCCTGTGGCGCTACATGTA 906

Qy        675 CACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTATTCA 734
                 ||||      |  || || |   |    || || ||||| ||||||||||| |
Db        907 TCACATGTACCAGCTGGACGCCTTCACACAATCGTGCCCGGCCGACCAGGACATTATCAA 966

Qy        735 TCACTATAAAGAACAAATGAATCTGTTCACAAATCAACGTGAAACCCTCCAATCGCCAAC 794
              |||||| ||    |||  || |||    |  ||  | |  ||    ||  |  |||| ||
Db        967 TCACTACAAGCTGCAACAGAGTCTCAAAATGAAGAAGCACGAGGAGCTGGAGACGCCCAC 1026

Qy        795 AAAAACGCACACAATTCC 812
                ||       |||||
Db       1027 GTTTACCACATACATTCC 1044


RESULT 3
AAC95377/c
ID   AAC95377 standard; cDNA; 786 BP.
XX
AC   AAC95377;
XX
DT   20-FEB-2001  (first entry)
XX
DE   Cat flea HMT Cl intracellular channel cDNA complement, SEQ ID NO:1876.
XX
KW   Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW   flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;
KW   detection; ss.
XX
OS   Ctenocephalides felis.
XX
PN   WO200061621-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US009437.
XX
PR   09-APR-1999;   99US-0128704P.
XX
PA   (HESK-) HESKA CORP.
XX
PI   Brandt KS,  Gaines PJ,  Stinchcomb DT,  Wisnewski N;
XX
DR   WPI; 2000-656323/63.
```

```
DR   P-PSDB; AAB29622.
XX
PT   Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT   acids useful for the prevention, diagnosis and treatment of flea
PT   infestations.
XX
PS   Claim 1; Page 901-902; 964pp; English.
XX
CC   The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC   acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC   or head and nerve cord (HNC) tissue. The invention also relates to the
CC   encoded proteins. The invention additionally encompasses expression
CC   constructs, recombinant viruses and recombinant cells comprising the
CC   nucleic acids of the invention, recombinant production of the proteins,
CC   antibodies against the proteins, a method of identifying inhibitors of
CC   the proteins, and compositions comprising the inhibitors for
CC   administration to an animal. The nucleic acids, and the proteins they
CC   encode may be used in the prevention, treatment and diagnosis of diseases
CC   associated with flea infestations. For example, the nucleic acids may be
CC   used to produce an HMT or HNC protein according to standard recombinant
CC   DNA methodology by inserting the nucleic acids into a host cell and
CC   culturing the cell to express the protein. The HMT and HNC nucleic acids
CC   may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC   and quantitate the presence of cat flea or other homologous nucleic acid
CC   sequences in samples. They may also be used to study the expression and
CC   function of the proteins and their role in metabolism. The HMT and HNC
CC   proteins may be used as antigens in the production of specific
CC   antibodies, and in assays to identify modulators (agonists and
CC   antagonists) of HMT and/or HNC protein expression and activity. The anti-
CC   HMT/HNC protein antibodies and antagonists may also be used to
CC   downregulate protein expression and activity. The antibodies may also be
CC   used as diagnostic agents for detecting the presence of flea polypeptides
CC   in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC   present sequence represents a cat flea HMT cDNA of the invention
XX
SQ   Sequence 786 BP; 182 A; 171 C; 178 G; 255 T; 0 U; 0 Other;

  Query Match             7.2%;  Score 63;  DB 3;  Length 786;
  Best Local Similarity   50.2%;  Pred. No. 3.3e-08;
  Matches  156;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy        462 ACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCT 521
              | ||  | | ||  |  ||||  |  |        || ||    || ||  |  |     |
Db        388 AAAGGATGACGTAAAAAGCAATAGTCTGCTGAGCCACCTGAGGAAAATCAACGACCATTT 329

Qy        522 ATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACT 581
              |   | |      |  |||  |||   ||  |||  |||      |  ||||| |||||
Db        328 GGCGCGGCGCGGCACGAGATTCCTAACGGGCGACACCATGTGCTGCTTCGACTGCGAACT 269

Qy        582 GATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCC 641
              ||||||  |  | || || ||  |  |    |    ||   |  | | ||| || |||||
Db        268 GATGCCCAGGTTACAACACATCAGGGTCGCCGCCAAGTATTTCGTCGAATTTGAAATTCC 209

Qy        642 ACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTAT 701
                  ||| | ||    | |||  ||||||          ||||    |  |||||  |
Db        208 GAGCAATCTAACCGCCTTATGGCGTTATATGTATCACATGTACCAGTTGGACGCATTCAC 149

Qy        702 TGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTT 761
              ||   || || ||||| || || || ||  | |||||||||   |||  | |   | |
Db        148 CCAGTCGTGCCCAGCCGACCAAGATATCATCAACCACTATAAACTGCAACAGCAGAGGAT 89

Qy        762 CACAAATCAAC 772
              ||  ||| | |
Db         88 CAGCAATAACC 78


RESULT 4
AAC95376
ID   AAC95376 standard; cDNA; 786 BP.
XX
AC   AAC95376;
XX
DT   19-FEB-2001  (first entry)

XX
DE   Cat flea HMT C1 intracellular channel cDNA ORF, SEQ ID NO:1875.
XX
KW   Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW   flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;
KW   detection; ss.
XX
OS   Ctenocephalides felis.
XX
PN   WO200061621-A2.
XX
PD   19-OCT-2000.
XX
PF   07-APR-2000; 2000WO-US009437.
XX
PR   09-APR-1999;   99US-0128704P.
XX
PA   (HESK-) HESKA CORP.
XX
PI   Brandt KS,  Gaines PJ,  Stinchcomb DT,  Wisnewski N;
XX
DR   WPI; 2000-656323/63.
DR   P-PSDB; AAB29622.
XX
PT   Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT   acids useful for the prevention, diagnosis and treatment of flea
PT   infestations.
XX
PS   Claim 1; Page 901; 964pp; English.
XX
CC   The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC   acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC   or head and nerve cord (HNC) tissue. The invention also relates to the
CC   encoded proteins. The invention additionally encompasses expression
CC   constructs, recombinant viruses and recombinant cells comprising the
CC   nucleic acids of the invention, recombinant production of the proteins,
CC   antibodies against the proteins, a method of identifying inhibitors of
CC   the proteins, and compositions comprising the inhibitors for
CC   administration to an animal. The nucleic acids, and the proteins they
CC   encode may be used in the prevention, treatment and diagnosis of diseases
CC   associated with flea infestations. For example, the nucleic acids may be
CC   used to produce an HMT or HNC protein according to standard recombinant
CC   DNA methodology by inserting the nucleic acids into a host cell and
CC   culturing the cell to express the protein. The HMT and HNC nucleic acids
CC   may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC   and quantitate the presence of cat flea or other homologous nucleic acid
CC   sequences in samples. They may also be used to study the expression and
CC   function of the proteins and their role in metabolism. The HMT and HNC
CC   proteins may be used as antigens in the production of specific
CC   antibodies, and in assays to identify modulators (agonists and
CC   antagonists) of HMT and/or HNC protein expression and activity. The anti-
CC   HMT/HNC protein antibodies and antagonists may also be used to
CC   downregulate protein expression and activity. The antibodies may also be
CC   used as diagnostic agents for detecting the presence of flea polypeptides
CC   in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC   present sequence represents a cat flea HMT cDNA of the invention
XX
SQ   Sequence 786 BP; 255 A; 178 C; 171 G; 182 T; 0 U; 0 Other;

  Query Match             7.2%;  Score 63;  DB 3;  Length 786;
  Best Local Similarity   50.2%;  Pred. No. 3.3e-08;
  Matches  156;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy        462 ACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCT 521
              | ||  | | ||  |  ||||  |  |        || ||    || ||  |  |     |
Db        399 AAAGGATGACGTAAAAAGCAATAGTCTGCTGAGCCACCTGAGGAAAATCAACGACCATTT 458

Qy        522 ATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACT 581
              |   | |      |  |||  |||   ||  |||  |||      |  ||||| |||||
Db        459 GGCGCGGCGCGGCACGAGATTCCTAACGGGCGACACCATGTGCTGCTTCGACTGCGAACT 518

Qy        582 GATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCC 641
              ||||||  |  | || || ||  |  |    |    ||   |  | | ||| || |||||
```

```
Db        519 GATGCCCAGGTTACAACACATCAGGGTCGCCGCCAAGTATTTCGTCGAATTTGAAATTCC 578

Qy        642 ACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTAT 701

Db        579 GAGCAATCTAACCGCCTTATGGCGTTATATGTATCACATGTACCAGTTGGACGCATTCAC 638

Qy        702 TGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTT 761

Db        639 CCAGTCGTGCCCAGCCGACCAAGATATCATCAACCACTATAAACTGCAACAGCAGAGGAT 698

Qy        762 CACAAATCAAC 772

Db        699 CAGCAATAACC 709
```

RESULT 5
AAC95375/c
ID AAC95375 standard; cDNA; 2383 BP.
XX
AC AAC95375;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea HMT Cl intracellular channel cDNA complement, SEQ ID NO:1874.
XX
KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;
KW detection; ss.
XX
OS Ctenocephalides felis.
XX
PN WO200061621-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US009437.
XX
PR 09-APR-1999; 99US-0128704P.
XX
PA (HESK-) HESKA CORP.
XX
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisnewski N;
XX
DR WPI; 2000-656323/63.
DR P-PSDB; AAB29622.
XX
PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT acids useful for the prevention, diagnosis and treatment of flea
PT infestations.
XX
PS Claim 1; Page 900-901; 964pp; English.
XX
CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC the proteins, and compositions comprising the inhibitors for
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The anti-
CC HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HMT cDNA of the invention
XX
SQ Sequence 2383 BP; 812 A; 374 C; 354 G; 835 T; 0 U; 8 Other;

```
  Query Match             7.2%;  Score 63;  DB 3;  Length 2383;
  Best Local Similarity   50.2%;  Pred. No. 5.5e-08;
  Matches  156;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy        462 ACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCT 521

Db       1926 AAAGGATGACGTAAAAAGCAATAGTCTGCTGAGCCACCTGAGGAAAATCAACGACCATTT 1867

Qy        522 ATCCGAGAGAAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACT 581

Db       1866 GGCGCGGCGCGGCACGAGATTCCTAACGGGCGACACCATGTGCTGCTTCGACTGCGAACT 1807

Qy        582 GATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCC 641

Db       1806 GATGCCCAGGTTACAACACATCAGGGTCGCCGCCAAGTATTTCGTCGAATTTGAAATTCC 1747

Qy        642 ACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTAT 701

Db       1746 GAGCAATCTAACCGCCTTATGGCGTTATATGTATCACATGTACCAGTTGGACGCATTCAC 1687

Qy        702 TGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTT 761

Db       1686 CCAGTCGTGCCCAGCCGACCAAGATATCATCAACCACTATAAACTGCAACAGCAGAGGAT 1627

Qy        762 CACAAATCAAC 772

Db       1626 CAGCAATAACC 1616
```

RESULT 6
AAC95374
ID AAC95374 standard; cDNA; 2383 BP.
XX
AC AAC95374;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea HMT Cl intracellular channel cDNA, SEQ ID NO:1872.
XX
KW Cat flea; hindgut and Malpighian tubule nucleic acid; HMT;
KW flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;
KW detection; ss.
XX
OS Ctenocephalides felis.
XX
PN WO200061621-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US009437.
XX
PR 09-APR-1999; 99US-0128704P.
XX
PA (HESK-) HESKA CORP.
XX
PI Brandt KS, Gaines PJ, Stinchcomb DT, Wisnewski N;
XX
DR WPI; 2000-656323/63.
DR P-PSDB; AAB29622.
XX
PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT acids useful for the prevention, diagnosis and treatment of flea
PT infestations.
XX
PS Claim 1; Page 896-898; 964pp; English.

XX
CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC the proteins, and compositions comprising the inhibitors for
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The anti-
CC HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HMT cDNA of the invention
XX
SQ Sequence 2383 BP; 835 A; 354 C; 374 G; 812 T; 0 U; 8 Other;

Query Match 7.2%; Score 63; DB 3; Length 2383;
Best Local Similarity 50.2%; Pred. No. 5.5e-08;
Matches 156; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 462 ACAGATTAAAGTTCACTACAATCGAGTCTGTGAGCAACTATCCAATATTGATCAGTTGCT 521
Db 458 AAAGGATGACGTAAAAAGCAATAGTCTGCTGAGCCACCTGAGGAAAATCAACGACCATTT 517

Qy 522 ATCCGAGAGAAAATCTCGATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACT 581
Db 518 GGCGCGGCGCGGCACGAGATTCCTAACGGGCGACACCATGTGCTGCTTCGACTGCGAACT 577

Qy 582 GATGCCACGTCTTCATCATATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCC 641
Db 578 GATGCCCAGGTTACAACACATCAGGGTCGCCGCCAAGTATTTCGTCGAATTTGAAATTCC 637

Qy 642 ACATAATTTCACTCATCTCTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTAT 701
Db 638 GAGCAATCTAACCGCCTTATGGCGTTATATGTATCACATGTACCAGTTGGACGCATTCAC 697

Qy 702 TGAGAGTTGTCCCGCCGATCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTT 761
Db 698 CCAGTCGTGCCCAGCCGACCAAGATATCATCAACCACTATAAACTGCAACAGCAGAGGAT 757

Qy 762 CACAAATCAAC 772
Db 758 CAGCAATAACC 768

RESULT 7
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.
XX
OS Oryza sativa.

XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX
DR WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 5.9%; Score 51.2; DB 8; Length 2000;
Best Local Similarity 10.4%; Pred. No. 0.00019;
Matches 70; Conservative 294; Mismatches 302; Indels 4; Gaps 1;

Qy 75 GTCAGGAATTGATGCTCGCCGCATTGGAGCCGATCTTTTCTGTCAGGAATTCTGGATGGA 134
Db 1030 GTGTGTAGGTGCTKMWRTYTRSMSWTYAMWKKYTKYMTAYSSTWKWYWAYKWRAYAWSRS 971

Qy 135 GTTGTATGCTCTTTATGAGATTGGAGTTGCACGAGTCGAAGTGAAGACTGTCAACGTGAA 194
Db 970 RKTWWCTGGKRMATYCGTKMWAAGRWRWRMAWCWYCCMWKWWKMTSCMWWKYWRTWSCWY 911

Qy 195 TTCTGAAGCATTTAAGAA----GAACTTTCTCGGAGCACAACCACCGATTATGATTGAAG 250
Db 910 TMWWGAMRYAYYAMRRRRWTYKWSWRRMYWTMTKWAWTWMTCMCMAKWYMATGWATWMWW 851

Qy 251 AGGAAAAAGAGCTGACATACACTGATAATCGAGAGATTGAAGGACGGATCTTTCATTTGG 310
Db 850 RYTMYTYCYAMTCAKCKYKMAMTKWWTTWACAWRATSWRWRAMAGMRWKRYKMKRAYWWR 791

Qy 311 CAAAGGAATTCAATGTTCCACTCTTTGAAAAGGATCCATCCGCTGAGAAGAGAATAGAGA 370
Db 790 WWRCWKAGWARWMKSRYRWKWKKYATRYYWKMWAMTWWWSWRRWKSYRMWSGMGRMRWSA 731

Qy 371 ACTTGTACAGGAACTTCAAACTGTTCCTGCGAGCAAAAGTAGAGTTCGATAAGGGAAAAA 430
Db 730 WRYCSRMKCAKTKYASSARWTKRAKRSYRYRRRWYWKRKGWTYRYRYWRSCRMTRARMSK 671

Qy 431 AGGAGCCATCGAGAGTTGAAGATCTTCCAGCACAGATTAAAGTTCACTACAATCGAGTCT 490
Db 670 RRKWAGASMKSCWMYWRGARSMWYSKYSCSAKCCKKTRYMTSSYMSTGMYGMYSSYKSMS 611

Qy 491 GTGAGCAACTATCCAATATTGATCAGTTGCTATCCGAGAGAAAATCTCGATATCTACTTG 550
Db 610 WTSKMSYMGKMTCTMYTSMKGSTRRSKMGRWSGMSRMYMRWWKKMRKRKYMRYMKWKCTW 551

Qy 551 GAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCATATTCGAATTA 610

```
Db        550 RRCMCYRWGYTMYTTSRSRMMYTGRYKARYTSKRRYMWYKYRKYCWYYYYGMYMKCSYMM 491

Qy        611 TTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCTCTGGGCTTATA 670

Db        490 RYGYCKACKKCCYAMCWKAAYSGMMMYWYRKYSKWMRMSTKYMWSMWYKKCRSMKYGAKG 431

Qy        671 TCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGATCAGGACATTA 730

Db        430 CYGCKMWTYCSYGYMKWYTYMGSYKYSRCYKYMRMYMYKGWMYMMYYSAYSSMMTWYYYY 371

Qy        731 TTCATCACTA 740

Db        370 AKYWKYWYKR 361
```

RESULT 8
ADM02307
ID ADM02307 standard; cDNA; 2238 BP.
XX
AC ADM02307;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human cDNA of the invention SEQ ID NO:992.
XX
KW ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
PN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-723558/69.
DR P-PSDB; ADM04750.
XX
PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
PS Claim 1; SEQ ID NO 992; 305pp; English.
XX
CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.
XX
SQ Sequence 2238 BP; 668 A; 440 C; 465 G; 665 T; 0 U; 0 Other;

Query Match 5.5%; Score 48; DB 11; Length 2238;
Best Local Similarity 50.4%; Pred. No. 0.0018;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599

Db        637 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 696

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659

Db        697 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 756

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719

Db        757 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 816

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771

Db        817 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 868
```

RESULT 9
ADA24507
ID ADA24507 standard; cDNA; 2887 BP.
XX
AC ADA24507;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human cDNA differential expressed in adipose tissue, INCYTE1383585.7.
XX
KW ss; differential expression; adipose tissue; cytostatic; hypotensive;
KW antiarteriosclerotic; antidiabetic; anorectic; gene therapy;
KW peroxisome proliferator-activated receptor gamma; PPARgamma;
KW diabetes mellitus; obesity; hypertension; atherosclerosis; breast cancer;
KW prostate cancer; colon cancer; polycystic ovarian syndrome.
XX
OS Homo sapiens.
XX
PN US2003096272-A1.
XX
PD 22-MAY-2003.
XX
PF 29-JUL-2002; 2002US-00208408.
XX
PR 30-JUL-2001; 2001US-0308868P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Schebye XM;
XX
DR WPI; 2003-606416/57.
XX
PT New combination comprising several cDNAs, useful for preparing a
PT composition for diagnosing or treating diabetes mellitus, obesity,
PT hypertension, atherosclerosis, or cancer of the breast, prostate or
PT colon.
XX
PS Claim 1; Page 47-48; 84pp; English.
XX
CC Then invention relates to a new combination comprising 55 cDNAs (ADA24485
CC -ADA24539)or their complements that are differentially regulated in an
CC adipose sample. Also included are detecting differential expression of
CC one or more cDNAs in a sample containing nucleic acids, screening several
CC molecules or compounds to identify a ligand that specifically binds a
CC cDNA, a vector comprising the cDNA, a host cell containing the vector,
CC producing a protein, screening several molecules or compounds, producing
CC an antibody and the isolated antibody. The cDNAs comprise sequences which
CC are upregulated or downregulated in response to peroxisome proliferator-
CC activated receptor gamma (PPARgamma) agonist. The combination comprising
CC several cDNAs is useful for preparing a composition for diagnosing or
CC treating diabetes mellitus, obesity, hypertension, atherosclerosis,
CC cancer of the breast, prostate or colon, or polycystic ovarian syndrome.
XX
SQ Sequence 2887 BP; 872 A; 552 C; 560 G; 903 T; 0 U; 0 Other;

Query Match 5.5%; Score 48; DB 9; Length 2887;
Best Local Similarity 50.4%; Pred. No. 0.002;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        816 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 875

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     | | |       |  ||     || |||||||| |  |  | |||    |
Db        876 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 935

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  || |   | |  |||||   ||
Db        936 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 995

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||       ||||  ||   |   ||     |||| ||  ||
Db        996 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 1047
```

RESULT 10
ACC83320
ID ACC83320 standard; DNA; 4290 BP.
XX
AC ACC83320;
XX
DT 22-AUG-2003 (first entry)
XX
DE Chloride intracellular channel 4 gene #SEQ ID 2.
XX
KW NF-ATc isoform a; chloride intracellular channel 4; osteopathic;
KW antiarthritic; antidiabetic; nephrotropic; osteoclast;
KW bone metabolic marker; arthritis; osteoporosis; arthritis deformans;
KW systemic erythematodex; bone reduction; diabetes; renal failure; gene;
KW ds.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   198. .959
FT   /*tag= a
FT   /product= "chloride intracellular channel 4"
XX
PN WO2003027284-A1.
XX
PD 03-APR-2003.
XX
PF 25-SEP-2002; 2002WO-JP009882.
XX
PR 26-SEP-2001; 2001JP-00294804.
PR 10-MAY-2002; 2002JP-00136030.
XX
PA (SUMU) SUMITOMO PHARM CO LTD.
PA (TAKE/) TAKEYA T.
XX
PI Takeya T, Hayashi K;
XX
DR WPI; 2003-393338/37.
DR P-PSDB; ABR62135.
XX
PT Bone metabolic disease marker and encoded polynucleotide for accelerated
PT differentiation/activation of osteoclasts, applicable in diagnosis of and
PT screening drugs for treating e.g. arthritis and osteoporosis.
XX
PS Claim 1; Page 61-63; 78pp; Japanese.
XX
CC The invention relates to a bone metabolic marker associated with the
CC accelerated differentiation or activation of osteoclasts. This consists
CC of a polynucleotide having not less than 15 consecutive bases in the base
CC sequence of the NF-ATc isoform a gene, or that of the chloride
CC intracellular channel 4 gene. The protein and its encoded polynucleotide
CC are applicable in diagnosis of and screening drugs for e.g. arthritis,
CC osteoporosis, arthritis deformans, systemic erythematodex, bone reduction
CC in diabetes and reduction of bone density in chronic renal failure. The
CC current sequence represents the chloride intracellular channel 4 gene
XX

SQ Sequence 4290 BP; 1287 A; 764 C; 832 G; 1407 T; 0 U; 0 Other;

  Query Match   5.5%; Score 48; DB 8; Length 4290;
  Best Local Similarity   50.4%; Pred. No. 0.0024;
  Matches 117; Conservative   0; Mismatches 115; Indels   0; Gaps   0;

```
Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        728 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 787

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     | | |       |  ||     || |||||||| |  |  | |||    |
Db        788 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 847

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  || |   | |  |||||   ||
Db        848 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 907

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||       ||||  ||   |   ||     |||| ||  ||
Db        908 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 959
```

RESULT 11
AAK51503
ID AAK51503 standard; cDNA; 4318 BP.
XX
AC AAK51503;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 48.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM78370.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 622-625; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The

```
CC   polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC   peptide therapy. The polypeptides have various cytokine-like activities,
CC   e.g. stem cell growth factor activity, haematopoiesis regulating
CC   activity, tissue growth factor activity, immunomodulatory activity and
CC   activin/inhibin activity and may be useful in the diagnosis and/or
CC   treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC   inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC   (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC   sequence listing were missing at the time of publication
XX
SQ   Sequence 4318 BP; 1319 A; 769 C; 814 G; 1416 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 48;  DB 4;  Length 4318;
  Best Local Similarity   50.4%;  Pred. No. 0.0024;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        697 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 756

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     |  | |       |  ||     || |||||||||| |  |  | |||    |
Db        757 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 816

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | ||||||| |||  |  |  || |   | |  ||||||   ||
Db        817 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 876

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||       ||||  ||  |   ||    ||||  ||  ||
Db        877 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 928


RESULT 12
AAI57882
ID   AAI57882 standard; cDNA; 4318 BP.
XX
AC   AAI57882;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polynucleotide SEQ ID NO 85.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US034263.
XX
PR   23-DEC-1999;   99US-00471275.
PR   21-JAN-2000; 2000US-00488725.
PR   25-APR-2000; 2000US-00552317.
PR   20-JUN-2000; 2000US-00598042.
PR   19-JUL-2000; 2000US-00620312.
PR   03-AUG-2000; 2000US-00653450.
PR   14-SEP-2000; 2000US-00662191.
PR   19-OCT-2000; 2000US-00693036.
PR   29-NOV-2000; 2000US-00727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y,  Qian XB,  Ren F,  Wang D;
PI   Wang J,  Wang Z,  Wehrman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J,  Zhao QA;
PI   Zhou P,  Goodrich R,  Drmanac RT;
XX
DR   WPI; 2001-442253/47.
DR   P-PSDB; AAM38726.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders such
PT   as central nervous system injuries.
XX
PS   Claim 1; SEQ ID NO 85; 10078pp; English.
XX
CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC   encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders. Note: The sequence data for this patent did not form
CC   part of the printed specification
XX
SQ   Sequence 4318 BP; 1319 A; 769 C; 814 G; 1416 T; 0 U; 0 Other;

  Query Match             5.5%;  Score 48;  DB 4;  Length 4318;
  Best Local Similarity   50.4%;  Pred. No. 0.0024;
  Matches  117;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        697 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 756

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |     |  | |       |  ||     || |||||||||| |  |  | |||    |
Db        757 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 816

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | ||||||| |||  |  |  || |   | |  ||||||   ||
Db        817 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 876

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||       ||||  ||  |   ||    ||||  ||  ||
Db        877 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 928


RESULT 13
ACF12844
ID   ACF12844 standard; cDNA; 4318 BP.
XX
AC   ACF12844;
XX
DT   10-SEP-2003  (first entry)
XX
DE   Human cervical cancer cell marker encoding cDNA SEQ ID NO:33.
XX
KW   Human; cervical cancer; cervical cancer marker; cancer therapy;
KW   detection; gene therapy; vaccine; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2002101075-A2.
XX
PD   19-DEC-2002.
XX
PF   12-JUN-2002; 2002WO-US018638.
XX
PR   13-JUN-2001; 2001US-0298155P.
PR   13-JUN-2001; 2001US-0298159P.
PR   14-NOV-2001; 2001US-0335936P.
XX
```

PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;
XX
DR WPI; 2003-156967/15.
DR P-PSDB; ABR92063.
XX
PT New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
PS Claim 4; Page 160-161; 386pp; English.
XX
CC ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials
XX
SQ Sequence 4318 BP; 1315 A; 764 C; 832 G; 1407 T; 0 U; 0 Other;

Query Match 5.5%; Score 48; DB 8; Length 4318;
Best Local Similarity 50.4%; Pred. No. 0.0024;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        728 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 787

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |      | | |        |  ||      || ||||||||| |  |  | |||    |
Db        788 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 847

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  || |  | |  ||||||   ||
Db        848 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 907

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||       ||||  ||   |   ||     |||| ||  ||
Db        908 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 959
```

RESULT 14
ADR25580
ID ADR25580 standard; DNA; 4318 BP.
XX
AC ADR25580;
XX
DT 21-OCT-2004 (first entry)
XX
DE Breast cancer prognosis marker #1441.
XX
KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004065545-A2.
XX
PD 05-AUG-2004.
XX
PF 15-JAN-2004; 2004WO-US001100.
XX
PR 15-JAN-2003; 2003US-00342887.
XX
PA (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
XX
DR WPI; 2004-593473/57.
XX
PT Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 1441; 226pp; English.
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
SQ Sequence 4318 BP; 1315 A; 764 C; 832 G; 1407 T; 0 U; 0 Other;

Query Match 5.5%; Score 48; DB 13; Length 4318;
Best Local Similarity 50.4%; Pred. No. 0.0024;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        728 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 787

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |      | | |        |  ||      || ||||||||| |  |  | |||    |
Db        788 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 847

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  || |  | |  ||||||   ||
Db        848 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 907

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||       ||||  ||   |   ||     |||| ||  ||
Db        908 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 959
```

RESULT 15
ADP23232
ID ADP23232 standard; cDNA; 4318 BP.
XX
AC ADP23232;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:326.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.

XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP23233.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 326; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 4318 BP; 1315 A; 764 C; 832 G; 1407 T; 0 U; 0 Other;

Query Match 5.5%; Score 48; DB 13; Length 4318;
Best Local Similarity 50.4%; Pred. No. 0.0024;
Matches 117; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

```
Qy        540 ATATCTACTTGGAAACAGTATGACTGAATATGACTGTGAACTGATGCCACGTCTTCATCA 599
              || |||   ||| ||    |||||   |  ||| ||  | ||| ||||    || |||
Db        728 ATTTCTGGATGGCAATGAAATGACATTAGCTGATTGCAACCTGCTGCCCAAACTGCATAT 787

Qy        600 TATTCGAATTATTGGATTGTCACTTCTTGGATTCGATATTCCACATAATTTCACTCATCT 659
              | |      |  | |        |  ||      || ||||||||| |   |  | |||    |
Db        788 TGTCAAGGTGGTGGCCAAAAAATATCGCAACTTTGATATTCCAAAAGAAATGACTGGCAT 847

Qy        660 CTGGGCTTATATCCTCACTGCATACCGTACAGCAGCATTTATTGAGAGTTGTCCCGCCGA 719
              ||||   ||  |    | |||||| |||  |  |  || |   | |  ||||||   ||
Db        848 CTGGAGATACCTAACTAATGCATACAGTAGGGACGAGTTCACCAATACCTGTCCCAGTGA 907

Qy        720 TCAGGACATTATTCATCACTATAAAGAACAAATGAATCTGTTCACAAATCAA 771
              | ||||  ||        ||||  ||   |   ||      |||| ||  ||
Db        908 TAAGGAGGTTGAAATAGCATATAGTGATGTAGCCAAAAGACTCACCAAGTAA 959
```

Search completed: March 10, 2005, 00:12:31
Job time : 595 secs